

# Designing a Data Management Plan: the GLOBAL-RURAL experience

---

PROFESSOR MICHAEL WOODS  
ABERYSTWYTH UNIVERSITY  
m.woods@aber.ac.uk



# GLOBAL-RURAL

---

ERC Advanced Grant

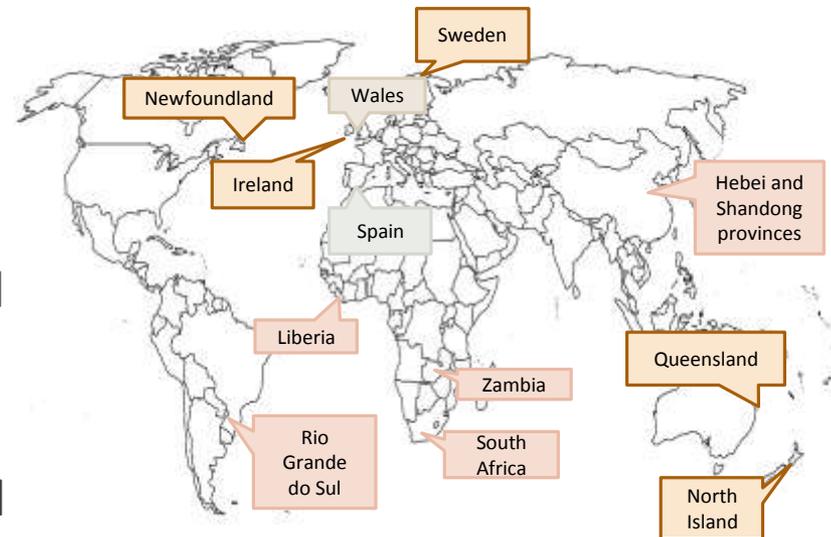
2014-2019

Examining how globalization is reproduced through rural localities and how rural agents respond to opportunities and challenges

9-strong team led by Professor Michael Woods at Aberystwyth University, UK

Research across 4 work packages and case studies in 12 countries

Mixed methods approach



Website and blog: [www.globalruralproject.wordpress.com](http://www.globalruralproject.wordpress.com)

# Types of Data

---

Publicly available secondary quantitative data (e.g. census statistics)

Commercially obtained secondary quantitative data

Newly generated quantitative data (e.g. survey data)

Interview data (recorded and transcribed)

Interview data (not recorded)

Focus group data

Observational / ethnographic notes

Archival data (historical documents, press articles etc)

Visual data (photographs, video)

Participant generated data (written, online etc)

# Data Sensitivities

---

Individual lifestyle data

Individual life course data

Business data

Information/opinion on  
controversial issues

Georeferenced data

Commercially purchased data

Data from semi-public spaces and  
forums

Third party data

Moving data internationally

Data collected in non-democratic  
states

# Context

---

Data protection and privacy laws in UK and case study countries

ERC policies

University policies

Research institute protocols

Gatekeeper/partners' policies

Guarantees made to research participants

Advice from ERC Ethics Advisor, University Ethics Advisor, University Data Protection Officer, Independent Ethics Advisor, colleagues and fellow researchers

# Context

Start of grant coincided with development of new Research Data Management policy by university



Appendix 3

**Manchester University (AU) Research Data Management (RDM) Service**

**1. Outline of RDM service available**

| Pre-requisites  | Throughout the project (To project)  | Post-project   |
|---|--|--|
| <ul style="list-style-type: none"> <li>• development with the preparation of data management plans, including:                             <ul style="list-style-type: none"> <li>- guidance on creating data management activities</li> <li>- the expert use of online tools.</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>• Review on data documentation, formats and standards to enable reuse</li> <li>• Guidance on coding, managing and archiving data to ensure compliance (e.g. physical and legal aspects)</li> <li>• Advise on any provision of access to data storage facilities that meet the needs of a wide range of data types, platforms and access needs.</li> </ul> | <ul style="list-style-type: none"> <li>• Advice on retaining data of long-term value</li> <li>• Support to make research data visible and/or available to defined audiences</li> <li>• Help for researchers in deciding how to archive data at the end of a project (or at any other appropriate point)</li> </ul> |

**2. RDM Service delivery**

- The priority outcomes of the RDM service to AU will be those creating data as part of funded research:
  - o BCUK
  - o Government
  - o European Commission
  - o European Research Council
  - o Charities
  - o Commercial
- A user project RDM service will be provided as a secondary priority to those who have produced data that underpins a research output and a storage solution is necessary for the data to be made available as part of the output's publication conditions
- The priority-owners for those delivering the RDM service will be data created by AU i.e. not third-party data but secondary data that has been derived from it.

Page 3 of 26

Appendix 3

**3. RDM Standards**

Where not explicitly provided by a funder the following standards are recommended within AU:

| Area                                       | Recommended standard  | Notes   |
|--|---|---|
| <b>RDM Service Usage</b>                   |   |   |
| <b>Data Management Plan (DMP) creation</b> | <p>Data Management Plans are required to be created for all new applications for funded research projects and should be included as an attachment with the submitted 'Study Plan'</p> <p>The Digital Location Centre's (DLC) DMP Online is recommended for use (<a href="https://dmponline.bcu.ac.uk/">https://dmponline.bcu.ac.uk/</a>)</p>                            | <ul style="list-style-type: none"> <li>• DMP Online provides templates for all major funders' Data Management' planning. Included Plans along with a general template where a specific one forms to suit work</li> <li>• All created plans can be exported to PDF format for attachment into applications within electronic submission systems</li> <li>• A general DMP system (the DMP) is available from: <a href="https://www.dcu.ie/~lib/dmp/Default.aspx?tabid=1030">https://www.dcu.ie/~lib/dmp/Default.aspx?tabid=1030</a></li> </ul>  |
| <b>Ethics</b>                              | <ul style="list-style-type: none"> <li>• If you are carrying out research involving human participants, consent forms must include a request for data to be shared and stored.</li> <li>• The methodology for anonymising data before sharing takes place should be decided within any accompanying participant information provided.</li> </ul>                        | <ul style="list-style-type: none"> <li>• The confidentiality of information supplied by research participants and the anonymity of respondents must be retained. However, where identifiers (names, addresses, contact details, images) can be removed from data before it is shared and 'redacted identifiers' (personal characteristics, workflow, occupation) can be aggregated.</li> <li>• See the UK Data Service's page on 'Anonymisation' for guidance: <a href="https://www.ukdataservice.ac.uk/datacatalogue/studies/study?id=10000">https://www.ukdataservice.ac.uk/datacatalogue/studies/study?id=10000</a></li> </ul> |
| <b>Legal compliance</b>                    | <ul style="list-style-type: none"> <li>• Subject to third party interests such as commercial partner/funder's research data which supports a scholarly work conducted by a member of University staff is owned by the University.</li> <li>• If third party data to be used within the project, all end user license terms and conditions should be kept and</li> </ul> | <p>For most partner projects, Intellectual Property rights (IPR) ownership should be agreed within a consortium agreement.</p>  |

Page 6 of 16

## Research at Aberystwyth

Research Excellence

Researcher Development

Research Good Practice

Research Ethics

Research Integrity Concordat

Open Access

▣ Research Data Management

What?

Why?

Who?

How?

Where?

Data Catalogue

Help

Working with Us

Research, Business & Innovation

Support for Researchers

Contact Us

Home » Research at Aberystwyth » Research Good Practice » Research Data Management

## Research Data Management (RDM)

### What?

What is RDM?

What is research data?

### Why?

Why should you manage your  
research data?

### Who?

Who requires RDM?  
Who can help you with RDM?

### When?

When should you manage your  
research data?

### How?

How to manage your research  
data  
How to share your research data

### Where?

Where can you store your data?  
Where to find data produced by  
Aberystwyth University

# Resources

---

<https://dmponline.dcc.ac.uk/>

The screenshot shows the homepage of the DMPonline website. At the top left is the DMPonline logo. To its right, a banner for an upcoming DCC workshop is displayed, dated 23 November 2015 in Glasgow. Below the logo is a navigation menu with links for Home, About, Roadmap, and Help. The main content area features a 'Welcome' message stating that DMPonline was developed by the Digital Curation Centre to help write data management plans. A 'Sign in' panel on the right contains input fields for email and password, a 'Remember me' checkbox, a 'Sign in' button, and a link for institutional credentials. Below the sign in panel is a 'Sign up' panel with a plus icon and a message for new users. A 'Screencast on how to use DMPonline' section is also visible, showing a thumbnail of a video player.

# Resources

Signed in as Michael Woods ▾



Upcoming DCC workshop - ['Advanced DMPonline: how to customise the tool for your institution'](#),  
23 November 2015, Glasgow

[View plans](#) [Create plan](#) [About](#) [Roadmap](#) [Help](#)

## Create a new plan

Please select from the following drop-downs so we can determine what questions and guidance should be displayed in your plan.

If you aren't responding to specific requirements from a funder or an institution, [select here to write a generic DMP](#) based on the most common themes.

**If applying for funding, select your research funder.**

Otherwise leave blank.

European Commission (Horizon 2020) ▾

[Not applicable/not listed.](#)

**To see institutional questions and/or guidance, select your organisation.**

You may leave blank or select a different organisation to your own.

Aberystwyth University ▾

[Not applicable/not listed.](#)

**Tick to select any other sources of guidance you wish to see.**

DCC guidance

# Resources

The templates provided for each phase are based on the annexes provided in the [Guidelines on Data Management in Horizon 2020](#) (4.16, 11 December 2019).

## Initial DMP

Export

Answer questions

An initial DMP should be completed within 6 months of starting the project.

The purpose of the Data Management Plan (DMP) is to provide an analysis of the main elements of the data management policy that will be used by the applicants with regard to all the datasets that will be generated by the project.

The DMP is not a fixed document, but evolves during the lifespan of the project.

The DMP should address the points below on a dataset by dataset basis and should reflect the current status of reflection within the consortium about the data that will be produced.

| Sections                                 | Questions  |
|--|--|
| For each data set specify the following: | <ul style="list-style-type: none"><li>- Data set reference and name</li><li>- Data set description</li><li>- Standards and metadata</li><li>- Data sharing</li><li>- Archiving and preservation (including storage and backup)</li></ul> |

# Resources

<http://www.dcc.ac.uk/resources/data-management-plans>



## Checklist for a Data Management Plan, v4.0

Please cite as: DCC. (2013). Checklist for a Data Management Plan, v4.0. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/data-management-plans>

| DCC Checklist                              | DCC Guidance and questions to consider   |
|--|--|
| <b>Administrative Data</b>                 |  |
| ID   | A pertinent ID as determined by the funder and/or institution.   |
| Funder                                     | State research funder if relevant.   |
| Grant Reference Number                     | Enter grant reference number if applicable (POST-AWARD DMPS ONLY).   |
| Project Name                               | If applying for funding, state the name exactly as in the grant proposal.  |
| Project Description                        | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- What is the nature of your research project?</li> <li>- What research questions are you addressing?</li> <li>- For what purpose are the data being collected or created?</li> </ul> <p><b>Guidance:</b></p> <p>Briefly summarise the type of study (or studies) to help others understand the purposes for which the data are being collected or created.</p>  |
| PI / Researcher                            | Name of Principal Investigator(s) or main researcher(s) on the project.  |
| PI / Researcher ID                         | E.g. ORCID <a href="http://orcid.org/">http://orcid.org/</a>   |
| Project Data Contact                       | Name (if different to above), telephone and email contact details.   |
| Date of First Version                      | Date the first version of the DMP was completed.   |
| Date of Last Update                        | Date the DMP was last changed.   |
| Related Policies                           | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Are there any existing procedures that you will base your approach on?</li> <li>- Does your department/group have data management guidelines?</li> <li>- Does your institution have a data protection or security policy that you will follow?</li> <li>- Does your institution have a Research Data Management (RDM) policy?</li> <li>- Does your funder have a Research Data Management policy?</li> <li>- Are there any formal standards that you will adopt?</li> </ul> <p><b>Guidance:</b></p> <p>Cite any other relevant funder, institutional, departmental or group policies on data management, data sharing and data security. Some of the information you give is the remainder of the DMP will be determined by the content of other policies. If so, point/link to them here.</p> |
| <b>Data Collection</b>                     |  |
| What data will you collect or create?      | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- What type, format and volume of data?</li> <li>- Do your chosen formats and software enable sharing and long-term access to the data?</li> <li>- Are there any existing data that you can reuse?</li> </ul> <p><b>Guidance:</b></p> <p>Give a brief description of the data, including any existing data or third party sources that will be used, in each case noting its content, type and coverage. Define and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.</p>  |
| How will the data be collected or created? | <p><b>Questions to Consider:</b></p> <ul style="list-style-type: none"> <li>- What standards or methodologies will you use?</li> <li>- How will you structure and name your folders and files?</li> <li>- How will you handle versioning?</li> <li>- What quality assurance processes will you adopt?</li> </ul> <p><b>Guidance:</b></p> <p>Outline how the data will be collected/created and which community data standards (if any) will be used. Consider how the data will be organised during the project, mentioning</p>  |

|  |  |
|--|--|
|  | For example naming conventions, version control and folder structures. Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.  |
| <b>Documentation and Metadata</b>  |  |
| What documentation and metadata will accompany the data?                     | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- What information is needed for the data to be read and interpreted in the future?</li> <li>- How will you capture / create this documentation and metadata?</li> <li>- What metadata standards will you use and why?</li> </ul> <p><b>Guidance:</b></p> <p>Describe the types of documentation that will accompany the data to help secondary users to understand and reuse it. This should at least include basic details that will help people to find the data, including who created or contributed to the data, its date of creation and under what conditions it can be accessed. Documentation may also include details on the methodology used, analytical and procedural information, definitions of variables, vocabularies, units of measurement, any assumptions made, and the format and file type of the data. Consider how you will capture this information and where it will be recorded. Wherever possible you should identify and use existing community standards.</p> |
| <b>Ethics and Legal Compliance</b>   |  |
| How will you manage any ethical issues?                                      | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Have you gained consent for data preservation and sharing?</li> <li>- How will you protect the identity of participants if required? e.g. via anonymisation</li> <li>- How will sensitive data be handled to ensure it is stored and transferred securely?</li> </ul> <p><b>Guidance:</b></p> <p>Ethical issues affect how you store data, who can see/use it and how long it is kept. Managing ethical concerns may include: anonymisation of data; referral to departmental or institutional ethics committees; and formal consent agreements. You should show that you are aware of any issues and have planned accordingly. If you are carrying out research involving human participants, you must also ensure that consent is requested to allow data to be shared and reused.</p>   |
| How will you manage copyright and Intellectual Property Rights (IPR) issues? | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Who owns the data?</li> <li>- How will the data be licensed for reuse?</li> <li>- Are there any restrictions on the reuse of third-party data?</li> <li>- Will data sharing be postponed / restricted e.g. to publish or seek patents?</li> </ul> <p><b>Guidance:</b></p> <p>State who will own the copyright and IPR of any data that you will collect or create, along with the funder(s) for its use and reuse. For multi-partner projects, IPR ownership may be worth covering in a consortium agreement. Consider any relevant funder, institutional, departmental or group policies on copyright or IPR. Also consider permissions to reuse third-party data and any restrictions needed on data sharing.</p>  |
| <b>Storage and Backup</b>  |  |
| How will the data be stored and backed up during the research?               | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Do you have sufficient storage or will you need to include charges for additional services?</li> <li>- How will the data be backed up?</li> <li>- Who will be responsible for backup and recovery?</li> <li>- How will the data be recovered in the event of an incident?</li> </ul> <p><b>Guidance:</b></p> <p>State how often the data will be backed up and to which locations. How many copies are being made? Storing data on laptops, computer hard drives or external storage devices alone is very risky. The use of robust, managed storage provided by university IT teams is preferable. Similarly, it is normally better to use automatic backup services provided by IT Services than rely on manual processes. If you choose to use a third-party service, you</p>   |

# Resources

<http://www.dcc.ac.uk/resources/data-management-plans>

|  |  |
|--|--|
|  | <p>should ensure that this does not conflict with any funder, institutional, departmental or group policies, for example in terms of the legal jurisdiction in which data are held or the protection of sensitive data.</p>  |
| How will you manage access and security?                 | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- What are the risks to data security and how will these be managed?</li> <li>- How will you control access to keep the data secure?</li> <li>- How will you ensure that collaborators can access your data securely?</li> <li>- If creating or collecting data in the field how will you ensure its safe transfer into your main secured systems?</li> </ul> <p><b>Guidance:</b></p> <p>If your data is confidential (e.g. personal data not already in the public domain, confidential information or trade secrets), you should outline any appropriate security measures and note any formal standards that you will comply with e.g. ISO 27001.</p>   |
| <b>Selection and Preservation</b>                        |  |
| Which data should be retained, shared, and/or preserved? | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- What data must be retained/destroyed for contractual, legal, or regulatory purposes?</li> <li>- How will you decide what other data to keep?</li> <li>- What are the foreseeable research uses for the data?</li> <li>- How long will the data be retained and preserved?</li> </ul> <p><b>Guidance:</b></p> <p>Consider how the data may be reused e.g. to validate your research findings, conduct new studies, or for teaching. Decide which data to keep and for how long. This could be based on any obligations to retain certain data, the potential reuse value, what is economically viable to keep, and any additional effort required to prepare the data for data sharing and preservation. Remember to consider any additional effort required to prepare the data for sharing and preservation, such as changing file formats.</p> |
| What is the long-term preservation plan for the dataset? | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Where e.g. in which repository or archive will the data be held?</li> <li>- What costs if any will your selected data repository or archive charge?</li> <li>- Have you costed in time and effort to prepare the data for sharing / preservation?</li> </ul> <p><b>Guidance:</b></p> <p>Consider how datasets that have long-term value will be preserved and curated beyond the lifetime of the grant. Also outline the plans for preparing and documenting data for sharing and archiving. If you do not propose to use an established repository, the data management plan should demonstrate that resources and systems will be in place to enable the data to be curated effectively beyond the lifetime of the grant.</p>  |
| <b>Data Sharing</b>                                      |  |
| How will you share the data?                             | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- How will potential users find out about your data?</li> <li>- With whom will you share the data, and under what conditions?</li> <li>- Will you share data via a repository, handle requests directly or use another mechanism?</li> <li>- When will you make the data available?</li> <li>- Will you pursue getting a persistent identifier for your data?</li> </ul> <p><b>Guidance:</b></p> <p>Consider where, how, and to whom data with acknowledged long-term value should be made available. The methods used to share data will be dependent on a number of factors such as the type, size, complexity and sensitivity of data. If possible, mention earlier examples to show a track record of effective data sharing. Consider how people might acknowledge the reuse of your data.</p>  |
| Are any restrictions on data sharing required?           | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- What action will you take to overcome or minimise restrictions?</li> <li>- For how long do you need exclusive use of the data and why?</li> <li>- Will a data sharing agreement (or equivalent) be required?</li> </ul> <p><b>Guidance:</b></p> <p>Outline any expected difficulties in sharing data with acknowledged long-term value.</p>  |

|   |   |
|---|---|
|   | <p>along with causes and possible measures to overcome these. Restrictions may be due to confidentiality, lack of consent agreements or IFR, for example. Consider whether a non-disclosure agreement would give sufficient protection for confidential data.</p>   |
| <b>Responsibilities and Resources</b>                 |   |
| Who will be responsible for data management?          | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Who is responsible for implementing the DMP, and ensuring it is reviewed and revised?</li> <li>- Who will be responsible for each data management activity?</li> <li>- How will responsibilities be split across partner sites in collaborative research projects?</li> <li>- Will data ownership and responsibilities for DMP be part of any consortium agreement or contract agreed between partners?</li> </ul> <p><b>Guidance:</b></p> <p>Outline the roles and responsibilities for all activities e.g. data capture, metadata production, data quality, storage and backup, data archiving &amp; data sharing. Consider who will be responsible for ensuring relevant policies will be respected. Individuals should be named where possible.</p> |
| What resources will you require to deliver your plan? | <p><b>Questions to consider:</b></p> <ul style="list-style-type: none"> <li>- Is additional specialist expertise (or training for existing staff) required?</li> <li>- Do you require hardware or software which is additional or exceptional to existing institutional provision?</li> <li>- Will charges be applied by data repositories?</li> </ul> <p><b>Guidance:</b></p> <p>Carefully consider any resources needed to deliver the plan, e.g. software, hardware, technical expertise, etc. Where dedicated resources are needed, these should be outlined and justified.</p>   |

# Key Issues

---

## Data Collection

What data will be collected?

Primary and secondary qualitative and quantitative data

Data collected primarily concerns organizations, relations, policies and events

Specified where individual personal data would be collected

How personal data might be used and conditions of use

# Key Issues

---

## Data Collection

How will data be collected?

Data collected through interviews, archival and library research and field observation

Some data collected through participatory methods, e.g. workshops, interactive walks, participatory GIS

Participation by informed consent

Specific consent obtained for photography involving individuals or on private property

# Key Issues

---

## Data Collection

How will data be recorded?

Interviews and workshops digitally recorded, with participant's consent, and professionally transcribed

If consent for recording not given, manual notes will be taken

Data recording in archives and libraries will follow institutional regulations and may involve photographing or scanning of documents and/or manual note-taking

# Key Issues

---

## Data Transfer

How will data be transferred from the field site?

Interview recordings and notes containing personal or sensitive information will be stored during fieldwork in password protected or encrypted files on a laptop or tablet computer and backed-up on a password-protected pen-drive or external hard drive

*Transfer of interview recordings for transcription?*

# Key Issues

---

## Data Processing

How will confidentiality and privacy be protected?

Personal data will be anonymized or disguised as soon as possible after collection

Use of pseudonyms

*Recording of appropriate metadata*

*Good practice in naming files*

*Delinking anonymized and non-anonymized data*

# Key Issues

---

## Data Storage

Where will data be stored?

Data, including interview transcripts and all associated documentation that could identify participants, will be stored on a secure drive on the *departmental server*

*Negotiation for appropriate secure driver space with university*

*University policy required data to be stored on a server physically located in the EU*

# Key Issues

---

## Data Storage

Who will have access?

Members of the GLOBAL-RURAL team

*Subsequently modified with permissions at folder level*

How will data be backed-up?

Automatic back-up on mirror server by university

Additional back-up of some data on password protected or encrypted pcs and external hard disk drives

# Key Issues

---

## Data Retention & Sharing

What data will be retained?

Anonymized processed data, for example in interview transcripts and databases.

Transcripts and other data where effective anonymization is not possible will be with-held from archiving

*Encouraged to think more expansively by Independent Ethics Advisor*

# Key Issues

---

## Data Retention & Sharing

Where will data be archived?

UK Data Archive

*But – evolving environment. Other possibilities?*

Who will have access to archived data?

Anonymized transcripts and other data will be made available for use by other researchers, except in cases where consent has been withheld by the participant

# Key Issues

---

## Data Retention & Sharing

What data will be destroyed and when?

Raw data from which research participants may be identifiable, including recordings of interviews, field notes etc, will be retained for the duration of the project and no more than 12 months after the end

These materials will be will destroyed no later than 12 months after the end of the project

*But – compliance with university policy?*

# Specific Issues

---

## **Data from 'non-conventional' methods**

Includes participatory methods, participant generated data (e.g. uploads to websites, social media data etc

Issues of consent, recording and anonymization, intellectual property and data ownership

## **Data from work with schools**

Compliance with two sets of policy

Procedures for informed consent

Intellectual property and data sharing

# Specific Issues

---

## **GIS analysis and mapping of data**

Appropriate georeferencing

Safeguards that locational data does not compromise anonymity

## **Data from research in ICPC countries (Brazil, China, Liberia, South Africa and Zambia)**

Attention to local legal regulations and cultural sensitivities

Advice and support from local partners

Role of field assistants and interpreters

Balancing security of data with local legal compliance

Exported data not to include any personal data or data with sensitive commercial or political applications

What data needs to be collected and what needs to be recorded?

# Reflections

---

Make use of tools and resources to support preparation of data management plans

Take advice widely and particularly on technical issues

Avoid compromising the potential of research for easy options

Look for creative solutions

A data management plan is an evolving framework

A good data management plan can improve the research process and outcomes