ESRC Template

Plan Name ESRC Template

Principal Investigator / **Researcher** TUoS Researcher

Plan Description -

Funder -

Institution University of Sheffield

Assessment of existing data

Provide an explanation of the existing data sources that will be used by the research project, with references

ESRC Guidance

When creating new data sources, explain why existing data sources can not be reused. If purchasing or reusing existing data sources, explain whether issues such as copyright and IPR have been addressed to ensure that the data can be shared i.e. explain how you plan to deal with permissions to share data you have created which is derived from data which you do not own.

Data sources that can be consulted are:

- Discover UK Data Service, with over 6,000 datasets of key economic, social and historical data spanning many disciplines and themes
- ESRC Research Catalogue of past and present research awards and their outputs
- RCUK Gateway to research of past and present research awards and their outputs

DCC guidance on Existing Data

Questions to consider:

- Are there any existing data or methods that you can reuse?
- Do you need to pay to reuse existing data?
- Are there any restrictions on the reuse of third-party data?
- Can the data that you create - which may be derived from third-party data - be shared?

Guidance:

Check to see if there are any existing data that you can reuse, for examples by consulting relevant repositories. When creating new data sources, explain why existing data sources cannot be reused. If purchasing or reusing existing data sources, explain how issues such as copyright and IPR have been addressed. A list of repositories is provided by Databib or Re3data.

The University of Sheffield: guidance on Existing Data

Please see the University of Sheffield webpage on 'Finding and reusing data' for guidance.

Provide an analysis of the gaps identified between the currently available and required data for the research

DCC guidance on Existing Data

Questions to consider:

- What is the relationship to existing data e.g. in public repositories?
- How does your data complement and integrate with existing data?

Guidance:

Consider the relationship between the data that you will capture and existing data available in public repositories or elsewhere.

The University of Sheffield: guidance on Existing Data

Please see the University of Sheffield webpage on 'Finding and reusing data' for guidance.

Information on new data

Provide information on the data that will be produced or accessed by the research project

ESRC Guidance

Give a brief description of new data which you envisage creating. This information should include how the data will be collected (in line with the proposed research methods), their format (e.g. SPSS, Open Document Format, tab-delimited format, MS Excel), and how they will be documented.

For example, cover:

- data volume
- data type
- data quality, formats, standards documentation and metadata
- methodologies for data collection and/or processing
- source and trustworthiness of third party data

Using standardised and interchangeable data formats ensures the long-term usability of data. Clear and detailed data descriptions and annotation, together with user-friendly accompanying documentation on methods and contextual information, makes data easy to understand and interpret and therefore shareable and with long-lasting usability.

- **Guidance on data formats**
- **Guidance on documenting data**

DCC guidance on Data Volumes

Questions to consider:

- Do you have sufficient storage?
- Do you need to include costs for additional managed storage?
- Will the scale of the data pose challenges when sharing or transferring data between sites?

Guidance:

Consider the implications of data volumes in terms of storage, backup and access. Estimate the volume of data in MB/GB/TB and how this will grow to make sure any additional storage and technical support required can be provided.

The University of Sheffield: guidance on Data Volumes

Please see the University of Sheffield Corporate Information and Computing Services webpages on 'Research data storage' and 'Storage options' for guidance.

DCC guidance on Data Type

Questions to consider:

- What types of data will you create?
- Which types of data will have long-term value?

Guidance:

Outline the types of data that are expected to be produced from the project e.g. quantitative, qualitative, survey data, experimental measurements, models, images, audiovisual data, samples etc. Include the raw data arising directly from the research, the reduced data derived from it, and published data.

The University of Sheffield: guidance on Data Type

Please see the University of Sheffield webpage 'What is research data?' for guidance.

DCC guidance on Data Format

Questions to consider:

- What format will your data be in?
- Why have you chosen to use particular formats?
- Do the chosen formats and software enable sharing and long-term validity of data?

Guidance:

Outline and justify your choice of format e.g. SPSS, Open Document Format, tab-delimited format, MS Excel. Decisions may be based on staff expertise, a preference

for open formats, the standards accepted by data centres or widespread usage within a given community. Using standardised and interchangeable or open lossless data formats ensures the long-term usability of data.

See UKDS Guidance on recommended formats.

The University of Sheffield: guidance on Data Format

Please see the University of Sheffield webpage on 'Organising your data: Choosing data formats' for guidance.

Quality assurance of data

Describe the procedures for quality assurance that will be carried out on the data collected at the time of data collection, data entry, digitisation and data checking.

ESRC Guidance

Quality control of data is an integral part of a research process. Describe the procedures for quality assurance that will be carried out on the data collected at the time of data collection, data entry, digitisation and data checking.

For example this might include:

- Documenting the calibration of instruments
- Taking duplicate samples or measurements
- Standardised data capture, data entry or recording methods
- Data entry validation techniques
- Methods of transcription
- Peer review of data
- Guidance on data quality control

DCC guidance on Data Quality

Questions to consider:

- How will you control data capture to ensure data quality?
- What quality assurance processes will you adopt?

Guidance:

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.

The University of Sheffield: guidance on Data Quality

Please see the UK Data Service webpage on data quality assurance for guidance.

Backup and security of data

Please describe the data back-up procedures that you will adopt to ensure the data and metadata are securely stored during the lifetime of the project.

ESRC Guidance

You may need to discuss your institution's policy on backups. If your data is sensitive (e.g. detailed personal data) you should discuss appropriate security measures which you will be taking.

The methods of version control of data files should also be stated. Version control includes making sure that if the information in one file is altered, the related information in other files is also adapted, as well as keeping track of versions of data files and their locations.

- Guidance on storing, backup and security
- Guidance on version control

DCC guidance on Data Security

Questions to consider:

- What are the risks to data security and how will these be managed?
- Will you follow any formal standards?

Guidance:

If your data is sensitive (e.g. detailed personal data, politically sensitive information or trade secrets) you should discuss any appropriate security measures that you will be taking. Note the main risks and how these will be managed. Identify any formal standards that you will comply with e.g. ISO 27001.

See DCC Briefing Paper on Information Security Management - ISO 27000.

See UKDS guidance on data security.

The University of Sheffield: guidance on Data Security

Note: Data security is needed to prevent unauthorised access or disclosure and changes to or destruction of data. Please see the University of Sheffield webpage 'Keeping your data safe' for guidance.

All staff and researchers must complete the online training at https://infosecurity.shef.ac.uk. This training will be particularly helpful when filling this part of the DMP. If you require assistance please contact the University's Information Security team at https://www.shef.ac.uk/cics/infosec.

The University has policies relating to information security requiring its users to adhere, as a minimum, to the following security standards: Information Security

Policy, Data Protection Policy. More secure system policies may be defined where necessary, for example where patient data is involved. University Departments may have established their own policies regarding information security, e.g. ScHARR Information Governance Policy.

The University of Sheffield is not an accredited ISO 27001 institution. The University provides email, Contacts and calendaring services, Google Drive and Google Sites through the Google Apps for Education suite. Google Apps for Education (and the data centres that support the service) are SSAE 16 / ISAE 3402 Type II SOC 2 audited and have achieved ISO 27001 certification. The University is satisfied that personal data is being processed appropriately in accordance with UK Data Protection Law and the University's own privacy policies, and that the security controls put in place by Google are sufficient to protect University data.

Example text for high-risk data: *"We will comply with the Data Protection Act and the University's own Information Security and Data Protection Policies. The project is governed by Department of Health and so we will comply with the NHS IGT."*

Example text for high-risk data: *"We recognize that this data is highly confidential and is critical to the clinical treatment of patients. Therefore a project specific security policy has been developed in conjunction with the University's Information Security Team [link to policy]".*

Suggested text for low-risk data: *"The data will will not include personal data relating to human participants. The University's Information Security Policies will be abided by at all times."*

Faculty of Medicine, Dentistry & Health: guidance on Data Security

Some Faculty of Medicine, Dentistry and Health Departments have established their own policies regarding information security, e.g. ScHARR Information Governance Policy.

DCC guidance on Storage and Backup

Questions to consider:

- Where will the data be stored?
- How will the data be backed up? i.e. how often, to where, how many copies, is this automated…
- Who will be responsible for storage and backup?
- Do you have access to enough storage or will you need to include charges for additional services?

Guidance:

Describe how the data will be stored and backed-up to ensure the data and metadata are securely stored during the lifetime of the project. Storing data on laptops, computer hard drives or external storage devices alone is very risky. The use of robust,

managed storage with automatic backup, for example that provided by university IT teams, is preferable.

See UKDA guidance on data storage and backup.

The University of Sheffield: guidance on Storage and Backup

Note: Storing data on laptops, computer hard drives or external storage devices alone is not recommended. The use of robust, managed storage with automatic backup is preferred by the University and by funders.

Data and definitive project documentation should be stored on centrally provisioned University of Sheffield virtual servers and research data storage infrastructure throughout the lifetime of the project. Both Windows and Linux Virtual Servers with up to 10TB of storage are made available to research projects. Access control is by authorised University computer account username and password. Off-site access is facilitated by secure VPN connection authenticated by University username and remote password. By default, two copies of data are kept across two physical plant rooms, with a 28 day snapshot made of data and backed up securely offsite at least daily. This service is maintained by the University's Corporate Information and Computing Services.

Google Drive may be used for more flexible collaborative working but only where non personal-sensitive information is involved. Where Google Drive is used, copies of complete and definitive documents should be transferred to the main project repository on the University research storage infrastructure.

Please see the University of Sheffield webpage on 'Keeping your data safe' for further guidance.

Faculty of Medicine, Dentistry & Health: guidance on Storage and Backup

All requests for research data storage in the Faculty of Medicine, Dentistry and Health should be made to the Faculty IT Hub in the first instance (med-it@sheffield.ac.uk). They will work with you to create an appropriate folder structure and give access to authorised users.

Management and curation of data

Outline your plans for preparing, organising and documenting data.

ESRC Guidance

Outline your plans for preparing and documenting data for sharing and archiving (unless otherwise agreed). Identify any additional plans for data sharing, if any. A crucial part of making data user-friendly, shareable and with long-lasting usability is to ensure they can be understood and interpreted by other users. This requires clear and detailed data description, annotation and contextual information, as well as good-structured and well-organised data files.

- Guidance on documenting data
- Guidance on transcribing qualitative data

- Guidance on organising data

DCC guidance on Preservation Plan
Questions to consider:

- What is the long-term preservation plan for the dataset? e.g. deposit in a data repository
- Will additional resources be needed to prepare data for deposit or meet charges from data repositories?

Guidance:

Researchers should 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.

The University of Sheffield: guidance on Preservation Plan

Note: For guidance see the University of Sheffield webpages on 'Preserving your data' and 'Data repositories'.

Long term preservation and access may be best managed by using a specialist data repository. Your funder may specify a data repository to use, such as UK Data Service ReShare. Alternatively, look in re3data.org to find an appropriate repository. If no suitable repository is available you may deposit data in ORDA, the University of Sheffield data repository. Alternatively, if you need to regulate users' access through 'Data sharing agreements', data may be retained in the University's research storage infrastructure and registered in ORDA.

Suggested text in all cases: *"Data will be archived in line with the University of Sheffield's Research Data Management Policy, which is a component of the University's Policy on Good R&I Practices (the 'GRIP' Policy)."*

Where data is in paper format: *"Data collected in paper form will be routinely digitised and the paper form disposed of / stored for at least 10 years at our universities in secured areas."*

For data deposited in external data repositories: *"Research data selected for long-term preservation and sharing will be deposited in [name of repository/weblink]. The [name of repository] is openly accessible and searchable and will guarantee preservation of these data for ten years or more. Metadata records describing these data will be created in ORDA, the University of Sheffield research data registry and repository"*

Where some research data are being deposited in ORDA: *"Data that are not deposited in [name of repository/weblink] will be deposited in ORDA, a repository and*

registry of research data produced at the University of Sheffield, which will preserve data for ten years or more."

Where data is deposited in ORDA only: *"Data selected for long-term preservation and sharing will be deposited in ORDA, a repository and registry of research data produced at the University of Sheffield, which will guarantee preservation for ten years or more."*

Where data is being retained locally, but not made 'openly' accessible: *"Data selected for long-term preservation and sharing will be stored on centrally provisioned University of Sheffield virtual servers and research storage infrastructure (https://www.sheffield.ac.uk/cics/research) for at least ten years. Records of these data will be published in ORDA, a registry of research data produced at the University of Sheffield."*

DCC guidance on Documentation

Questions to consider:

- What metadata, documentation or other supporting material should accompany the data for it to be interpreted correctly?
- What information needs to be retained to enable the data to be read and interpreted in the future?

Guidance:

Describe the types of documentation that will accompany the data to provide secondary users with any necessary details to prevent misuse, misinterpretation or confusion. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data.

The University of Sheffield: guidance on Documentation

Note: Documentation and metadata describe the context, content and structure of your data and are essential for understanding and reusing them. See the University of Sheffield webapge 'Describing your data' for more information.

Example text: *"Methods and SOPs will be stored electronically in Microsoft Word documents (.doc) with the spreadsheets containing data"*

"Explanation of the experimental and analytical methods used will be provided in text documents, stored alongside the data"

"Data documentation will accompany datasets submitted to the ... repository at the end of the research"

Difficulties in data sharing and measures to overcome these

If you expect obstacles to sharing your data, explain which and the possible measures you can apply to overcome these.

ESRC Guidance

State explicitly which data may be difficult to share and why. If ethical issues could cause difficulties in data sharing, explain your strategies for dealing with these issues in the relevant section of the Je-S form, e.g. discussing data sharing with interviewees as part of consent discussions or anonymising data. Refer to the requirements of the ESRC Framework for Research Ethics.

The ESRC supports the position that most data - even sensitive and confidential data - can be curated and shared ethically and legally provided researchers pay attention right from the planning stages of research to the following aspects:

- when gaining informed consent, include consent for data sharing
- where needed, protect participants' identities by anonymising data; and
- address access restrictions to data in the data management and sharing plan, before commencing research.

Guidance on consent and ethics

Guidance on legislation relating to research data, including DPA and FoI

DCC guidance on Restrictions on Sharing

Questions to consider:

- Are any restrictions on data sharing required? e.g. limits on who can use the data, when and for what purpose.
- What restrictions are needed and why?
- What action will you take to overcome or minimise restrictions?

Guidance:

Outline any expected difficulties in data sharing, along with causes and possible measures to overcome these. Restrictions to data sharing may be due to participant confidentiality, consent agreements or IPR. Strategies to limit restrictions may include: anonymising or aggregating data; gaining participant consent for data sharing; gaining copyright permissions; and agreeing a limited embargo period.

The University of Sheffield: guidance on Restrictions on Sharing

Note: At the end of your research project, your funder may require you to make your research data available for sharing with as few restrictions as possible. Restrictions on the release of data may be allowed, to protect confidentiality and for other ethical and legal considerations:-

- Does your data include confidential and sensitive information?
- Have participants given consent for their data being shared?
- Consider what can be done to make sensitive data openly sharable - can these data be anonymised?
- If different parts of your research data require different access conditions, separate them and deposit them separately, applying different access conditions.

See the University of Sheffield 'Regulatory requirements' webpage and the UK Data Service 'Legal and ethical issues' webpages for more information.

Suggested text if no restrictions are foreseen: *"At present we do not foresee any delays in data sharing following publication of the main research findings."*

Suggested text for patient-based studies: *"Patients will be made aware of our data sharing procedures at the time of consent."*

DCC guidance on Managed Access Procedures

Questions to consider:

- Will access be tightly controlled or restricted? e.g. by using data enclaves / secure data services
- Will a data sharing agreement be required?
- How will the data be licensed for reuse?

Guidance:

Indicate whether external users will be bound by data sharing agreements, licenses or end-user agreements. If so, set out the terms and key responsibilities to be followed. Note how access will be controlled, for example by the use of specialist services. A data enclave provides a controlled secure environment in which eligible researchers can perform analyses using restricted data resources. Where a managed access process is required, the procedure should be clearly described and transparent.

The University of Sheffield: guidance on Managed Access Procedures

Note: Restrictions on the release of data may be allowed, to protect confidentiality and for other ethical and legal considerations. Access to and use of Sensitive and confidential data can be restricted and regulated using end user licenses, data sharing agreements or by using a data enclaves.

Whatever form of publishing is used, research data needs to be licensed to indicate what users may or may not do with the data. Data repositories will indicate what licenses are available for the data they house. More information is available from the Digital Curation Centre webpage 'How to license research data'.

See the University of Sheffield 'Regulatory requirements' webpage and the UK Data Service 'Legal and ethical issues' webpages for more information.

Suggested text: *"The University of Sheffield's Good Research and Innovation Practice (GRIP) Policy follows RCUK principles for data sharing (http://www.rcuk.ac.uk/research/datapolicy/)"*

Suggested text where a Data sharing agreement is required: *"External users will be bound by data sharing agreements as specified by the [name of funder] Data Sharing Policy."*

Suggested text where an external collaborator is involved: *"Data sharing agreements will be put in place with [name of collaborator], who will be a primary re-*

user of data"

Faculty of Medicine, Dentistry & Health: guidance on Managed Access Procedures

Access to and use of Sensitive and confidential data can be restricted and regulated using end user licenses or data sharing agreements.

MRC, Wellcome Trust and Cancer Research UK provide guidance on data sharing agreements in 'Good Practice Principles for Sharing Individual Participant Data from Publicly Funded Clinical Trials'.

MRC provides information on data access governance, facilitating data access and data sharing agreements in 'MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies' and 'Principles for access to, and use of, MRC funded research data'.

The University of Sheffield ScHARR Information Governance Policy provides guidance on information sharing and an example data sharing agreement.

Consent, anonymisation and strategies to enable further re-use of data

Make explicit mention of the planned procedures to handle consent for data sharing for data obtained from human participants, and/or how to anonymise data, to make sure that data can be made available and accessible for future scientific research.

ESRC Guidance

If unsure how issues of confidentiality are to be addressed to facilitate data sharing, get in touch for advice.

Guidance on consent and ethics

Guidance on anonymising data

DCC guidance on Ethical Issues

Questions to consider:

- Have you gained consent for data preservation and sharing?
- How will sensitive data be handled to ensure it is stored and transferred securely?
- How will you protect the identity of participants? e.g. via anonymisation or using managed access procedures

Guidance:

Investigators carrying out research involving human participants must ensure that consent is obtained to share data. Managing ethical concerns may include: anonymisation of data; referral to departmental or institutional ethics committees; and formal consent agreements. Ethical issues may affect how you store data, who can see/use it and how long it is kept. You should show that you're aware of this and have planned accordingly.

See UKDS guidance on consent for data sharing

The University of Sheffield: guidance on Ethical Issues

For guidance on ethical and legal issues please see the University of Sheffield webpage on 'Regulatory requirements'.

See also the UK Data Service webpages on 'Legal and ethical issues'.

Copyright and intellectual property ownership

Please state who will own the copyright and IPR of any new data that you will generate.

ESRC Guidance

Guidance on data copyright

DCC guidance on IPR Ownership and Licencing

Questions to consider:

- Who owns the data?
- How will the data be licensed for reuse?
- If you are using third-party data, how do the permissions you have been granted affect licensing?
- Will data sharing be postponed / restricted e.g. to seek patents?

Guidance:

State who will own the copyright and IPR of any new data that you will generate. For multi-partner projects, IPR ownership may be worth covering in a consortium agreement. If purchasing or reusing existing data sources, consider how the permissions granted to you affect licensing decisions. Outline any restrictions needed on data sharing e.g. to protect proprietary or patentable data.

See the DCC guide: How to license research data.

The University of Sheffield: guidance on IPR Ownership and Licencing

Please see the University of Sheffield webpage on 'Regulatory requirements' for guidance on IPR and data licensing.

The UK Data Service provides guidance on 'Rights relating to research data' including licensing.

The Digital Curation Centre provides guidance on 'How to license research data'.

Responsibilities

Outline responsibilities for data management within research teams at all partner institutions

ESRC Guidance

Indicate who within your research team will be responsible for data management,

metadata production, dealing with quality issues and the final delivery of data for sharing or archiving. Provide this information within the Staff Duties section in the Je-S form and where appropriate in the Justification of Resources. If several people will be responsible state their roles and responsibilities in the relevant section of the Je-S form. For collaborative projects explain the coordination of data management responsibilities across partners in your Data Management Plan.

- Guidance on data management roles and responsibilities
- Guidance on how to cost data management

DCC guidance on Responsibilities

Questions to consider:

- Who is responsible for each data management activity?
- How are responsibilities split across partner sites in collaborative research projects?

Guidance:

Outline the roles and responsibilities for all activities e.g. data capture, metadata production, data quality, storage and backup, data archiving & data sharing. Individuals should be named where possible. For collaborative projects you should explain the co-ordination of data management responsibilities across partners.

See UKDS guidance on data management roles and responsibilities.