

CIS Facility: Recommendations for data sharing

Through data sharing, researchers can make the best use of new and existing data, increasing the cost effectiveness of data collection and improving policy decision-making outcomes. Data sharing enables researchers to validate one another's research outcomes and allows for the combination of data from multiple sources across states and organisations, encouraging novel and diverse explorations of data.

The purpose of this document is to guide researchers in identifying strengths of and barriers to research collaboration, to better facilitate efficient, effective and secure data sharing. These considerations will form the basis of a mutually beneficial data sharing agreement.

Broadly, these considerations include; ensuring all partners have similar expectations of the data sharing objectives in terms of scope, governance and responsibilities, and that all barriers, ethical and confidentiality concerns have been considered, and there is agreement in how outcomes, including publications, will be managed. The data to be shared and the conditions under which the sharing will occur should be clearly outlined in a data agreement plan to ensure appropriate management and maintenance of the shared data.

1. Parameters of the data sharing request: purpose and scope and use of data

Collaborators need to approach sharing with a specific research question in mind, outlining how the data will be used, the analyses that will be performed, and the projected outcomes. In return, the custodians of the data should articulate any restrictions on how the data can be used; in particular, how the receiver may access, copy, share, transcribe, publish or disseminate data, derivative work products and research findings, and how ownership of any publications/reports or datasets arising from the data sharing will be determined. If the receiver intends to profit from the data findings or share data findings with a third party, this must be also discussed.

2. Data sharing processes: consent, confidentiality and data protection

It is the responsibility of the researchers to ensure that data sharing is undertaken legally, fairly, and within boundaries of the consent obtained from study participants and defined by existing privacy statements, protocols, procedures and statements. The use of data, in particular identifiable data, needs to be either covered by study consent contracts and ethics agreements or sufficiently anonymised as to allow data to be shared without explicit consent.

Any data sharing agreement should articulate protocols to prevent unpermitted use of shared data or disclosure of findings. The University HRECs will support researchers in determining if their proposed data sharing project comply with funder guidelines and requirements, publishers' data policies and human ethics requirements.

Measures taken to protect the data from unauthorised access, amendment, use or disclosure should be identified in a data sharing agreement. These may include explicit data storage, use, modification and record keeping protocols, licencing, data encryption, and confidentiality agreements for data handlers.

3. Monitoring of the data sharing agreements: governance and project management

Clearly defined project governance arrangements, including the roles and responsibilities for all involved researchers and timelines for the proposed project will improve the operational management of the data sharing agreement. The review, monitoring and amendment of the agreement should occur as required.

Individuals and organisations sharing data must have agreed upon roles and responsibilities regarding data access and management, governance, monitoring of procedures for data access, documentation and disclosures, overseeing performance milestones, and handling dispute resolution processes.

4. Publishing protocols: reporting, publication and attribution

A data sharing agreement should consider the expectations of data custodians and receivers regarding reporting, publication and attribution. As individual researchers may be hesitant to dedicate the time to preparing and curating data for sharing if there is no clear benefit, ensuring researchers are rewarded for their intellectual labour through appropriate attribution will make investing time in data sharing more attractive.

The nominated process for attribution, and agreements between researchers about what findings are made public, where and when these findings may be reported should be determined prior to data sharing. Collaborating researchers must comply with all prior publishing agreements to ensure data is reported appropriately, with correct attribution and approval from the authorised researchers, and within an acceptable timeframe. In reporting the findings of analyses derived from shared data, researchers must comply with all ethical agreements, regulatory authorities, and confidentiality/ consent agreements associated with the shared data.

5. Complaints, breaches and withdrawal from the data sharing agreement

The process for managing concerns or grievances will vary between research projects, depending on the sensitivity of the data involved. It may be necessary to ensure there are clear guidelines for compensation or other consequences for violating the data-sharing agreement.

The grounds by which the data sharing agreement may be terminated should be considered prior to data sharing. Any alleged breaches of the terms of the agreement should be discussed with all collaborating researchers with all participants given an opportunity to explain and amend on mutually agreeable terms via a predetermined dispute resolution process.

6. Financial costs of data-sharing

It may be necessary to clarify who will cover any monetary costs of sharing the data, if costs are incurred. The provider and the receiver of the data may share the costs, or it may be necessary for the receiver alone to source the funding required to cover data-sharing expenses.

Appendix A

Checklist of discussion points for a data sharing agreement:

		Discussed? (signed/date)
Parameters of the data sharing request		
1.1 Purpose and scope		
	What are the key research questions the data will address?	
	What is the planned end-use of the data?	
	What variables are being requested?	
	Will the data be combined or linked with any other data?	
	Is there enough information available to assess the data's fitness for purpose?	
1.2 Data to be shared		
	Is the data in a form appropriate for sharing? (final or subject to revision?)	
	Are the standards/classifications used in data generation clear?	
	What methodology was used to create or generate the data?	
	What descriptive or definitional metadata is available for the data?	
	Are technical and explanatory notes available for the data?	
	In what format will the data be shared?	
	What variables will be shared?	
	How often will the data be shared (if more than once)?	
	What processes are used to monitor and assess data quality?	
1.3 Intended use of the data		
	For what purpose has the data been requested?	
	How will the data be used by the data requestor?	
	How will the data be interpreted and/or analysed?	
1.4 Constraints on the use of the data		
	At this time, are there any known constraints on the use of the data? Including: restrictions on how the data requester may access, copy, share, transcribe, publish or disseminate data, derivative work products and research findings	
	Do the receivers intend to profit from the data findings and is this permitted?	
	Will the original data or derivative data sets be shared with any third parties?	
1.5 Authorised research partners		
	Who will have access to the shared data?	
	Has a data custodian been nominated?	
	Is there a protocol for allowing new researchers to gain access to the data?	
1.6 Governance of the proposed project data outcomes		
	Is it necessary to nominate additional governance roles? (for: accountability, monitoring procedures for accessing and documenting requests and disclosures, reviewing the progress on performance criteria, and dispute resolution processes)	
	By what process will these additional roles/responsibilities be distributed and revised?	
	Who will own the intellectual property developed during the project?	
	If data sets are to be retained after the project is completed, who will take custodianship of any new datasets created?	

Data sharing processes (mitigate risks and reduce barriers)		
2.1 Consent agreements and confidentiality		
	Is it possible for the requested data to be shared, within boundaries of the consent obtained from study participants?	
	Will the data need to be anonymized or de-identified prior to data sharing?	
	Is there a need to seek further consent from study participants prior to data sharing?	
	What safeguards are required to protect sensitive/identifying information?	
	Are there any addition barriers within the authorising environment, funding agreements, or legislation that may interfere with the sharing of data?	
	Has advice from the UoA HREC been sought for the proposed data sharing project?	
2.2 Measures to ensure adequate data protection		
	Have the following data protection measures been considered/developed?	
	<ul style="list-style-type: none"> • Protocols on how data will be stored (form, password protection) and accessed (physical, electronic) 	
	<ul style="list-style-type: none"> • Licensing 	
	<ul style="list-style-type: none"> • The need for data encryption 	
	<ul style="list-style-type: none"> • Procedures to allow non-authorized researcher access to data 	
	<ul style="list-style-type: none"> • Data quality control standards, data transcription standards 	
	<ul style="list-style-type: none"> • Confidentiality agreements for data handlers 	
	<ul style="list-style-type: none"> • Protocols on data saving, downloading, formatting, storage and backup 	
	<ul style="list-style-type: none"> • Agreements about modifications to the data and databases 	
	Have potential or actual risks to data security been identified and mitigation strategies developed?	
	Will further training be necessary for researchers who will have access to the data?	
	Will all authorized researchers have equal access to the data, or will access be restricted for some researchers?	
3. Monitoring and review of the data sharing agreements		
3.1 Governance of the data sharing protocol agreements		
	How will governance for the project been arranged?	
	Will governance roles be attributed to authorized researchers?	
	Will there be a plan for the review, monitoring and amendment of the data sharing agreement?	
	Will there be a formal review of the data sharing agreement? When will this review(s) occur?	
	Which responsibilities will be managed by a nominated governance officer (Table 5)?	
3.2 Management of project timelines		
	When will data and supporting documents will be made available to the receiver and how long the receiver will be able to use the data?	
	When will data sharing agreements will be formally reviewed, renewed or terminated?	
	What will be the retention periods for the data and a timeline for data deletion/destruction (with processes to ensure secure deletion takes place)?	
	What is the timeline for data analysis, reporting and projected publication?	

4. Attribution and publishing protocols		
4.1 Reporting		
	Are there any ethical agreements, regulatory authorities, and confidentiality/consent agreements associated with the shared data that may restrict reporting/publication?	
	Will there be protocols/governance arrangements regarding the reporting of data findings?	
4.2 Publication		
	Will the project have an official Publication Plan?	
	Who will be responsible for the development and maintenance of the plan?	
4.3 Attribution		
	What should be included to be considered an appropriate data citation?	
5. Complaints, breaches and withdrawal from the data sharing agreement		
5.1 Concerns/Grievances and breaches to the agreement		
	How will grievances/concerns about compliance with the data sharing agreement be raised?	
	How will a dispute resolution process be managed?	
	What grounds would be considered violation of the data sharing agreement?	
	What consequences are appropriate for violating the data-sharing agreement?	
5.2 Withdrawal		
	What grounds would be considered reasonable for allowing termination of the data sharing agreement?	
	What procedures must be followed in the circumstances of termination of the agreement?	
6. Other		
6.1 Financial costs of data-sharing		
	Are there any expected/foreseen monetary costs for this project?	
	Who will cover any monetary costs incurred from sharing the data?	

Table 1: Purpose and scope of the proposed project

Project code	<i>Unique identifier for the information</i>
Title	<i>title for the proposed data sharing project</i>
CI	<i>Names and locations of the CIs for this data sharing project</i>
Project goal	<i>The project goal that the information supports. Why is the information being shared?</i>
Source	<i>Data to be shared – enough details about the data set that it is recognisable (i.e. the name of study, database, spreadsheet), names/details of specific variables to be provided where appropriate</i>
Nominated data custodian	<i>The person who can provide expertise / advice about the information required. This person(s) will be a point of contact for clarification on the data set completeness/accuracy and should be familiar with the data but may not be the CI who collected the data.</i>
Description	<i>A brief description of the information to be shared – supported by Table 2</i>
Consent requirements	<i>Details of the ethics/consent agreements used to support the sharing or collection of the information, including if additional consent from participants needs to be requested</i>
Ethics requirements/details	<i>Details of ethics agreements for the data sharing project</i>
Legal restrictions	<i>Any legislative and consent based restrictions should be documented in a data sharing agreement</i>
Relevant protocols and agreements	<i>Details of any relevant information sharing protocols and agreements relating to this information within the context of this project, including data agreements that restrict the use and processing of the information -How data will be shared/stored, clear roles and responsibilities, timelines</i>
Acceptance criteria	<i>Measures that will determine if the requirements for data sharing have been met, including methodology/results quality standards</i>
Risks	<i>Any foreseeable risks associated with holding and sharing the information</i>
Monitoring procedures	<i>How the gathering and management of this information will be monitored?</i>
Costs	<i>Any foreseen financial costs and how they will be met</i>

Table 2: Data to be shared

Purpose	Data details	Data provider	Recipients
<i>State the function for which information is required</i>	<i>Eg. Variable names, supporting documents, data dictionaries</i>	<i>Researcher(s) who will provide the information</i>	<i>Researcher(s) who will receive and use the information</i>

Table 3: Authorised research partners

Name of researcher	Role in the proposed project	Responsibilities in the proposed project	Organisation/contact details	Date

Table 4: Measures to ensure adequate data protection

Data custodian(s)	<i>Name of the authorised researcher responsible for data custodianship (may be advisable to have two nominated persons for each side, the providers of the data and the receivers of the data)</i>
Data access procedures	<i>Step-by-step outline of how data will be delivered, managed</i>
Confidentiality	<i>Process for ensuring participant confidentiality</i>
Deletion procedures	<i>Timeline and protocol for deleting data after project completion</i>

Table 5: Governance roles

Role	Nominated officer	Date
Authorising the addition of new researchers for the project		
Amendments to the protocol will be reported to and signed off by the nominated researchers who will be responsible for endorsing changes to the protocol.		
Overview of authorising environment, legal authorities, obligations and prohibitions on disclosing information		
Compliance with agreed standards around how information is formatted and stored		
Development of guidelines for rules and responsibilities of data custodians when maintaining data sets, and the monitoring of access logs and performance criteria		
Monitoring of processes for assessing disclosure requests and procedures for documenting requests and disclosures		
Identification of potential/actual risks along with mitigation strategies		
Determining agreed common retention periods for the data and that processes to ensure secure deletion occur on schedule		