

Wellcome Rubric v2.1

Funder Template	Wellcome Trust			
Purpose of rubric	Providing feedback to researchers and guidance to peer-reviewers			
Notes	<p>A data management and sharing plan allows data creators to more effectively organise and document stages in their research.</p> <p>This includes:</p> <ul style="list-style-type: none"> - all proposals where the main goal is to create a database resource - research that generates significant datasets that could be shared, eg where the data could be used to address research questions other than those it was intended for. - a data sharing plan is compulsory for applications to Biomedical Science, Innovations and Humanities and Social Science funding schemes, if the data outputs are likely to be of value to other researchers and users. 			
Documents Used	<p>Wellcome Trust policy on data management and sharing (https://wellcome.ac.uk/funding/managing-grant/policy-data-management-and-sharing)</p> <p>Wellcome Trust data management plan guidance (https://wellcome.ac.uk/funding/managing-grant/developing-data-management-and-sharing-plan)</p>			
Useful resources	<p>Wellcome Trust advice on the choice of repositories (https://wellcome.ac.uk/funding/managing-grant/data-repositories-and-database-resources)</p> <p>Biosharing: searchable database of standards used in life sciences research (https://biosharing.org/)</p> <p>The Expert Advisory Group on Data Access (https://wellcome.ac.uk/what-we-do/our-work/expert-advisory-group-data-access)</p>			
Version History	<p>V0 - Performance criteria developed</p> <p>V1.0 - Performance levels developed</p> <p>v1.1 - Feedback from other members of project incorporated</p> <p>v1.2 - Feedback from funder incorporated (22Jan17)</p> <p>v2.0 - Copy made to upload to Research Data Network (23Jan17).</p>			
	Performance Criteria		Performance Levels	
		Detailed	Addressed but incomplete / unsatisfactory	Not addressed
	What types of data outputs will your research generate?	Data types clearly defined. Eg experimental measurements, models, recordings, video, images, machine logs, source code, databases, physical samples etc.	Some data types are mentioned, but not all.	No information provided.
	Which data will have value to other people and why?	Data types of potential value to others clearly identified and justification about the value is provided (indication of likely user base/demand).	Valuable data types merely listed, but no justification of the value provided.	No information provided.
	Will file formats in which data will be stored and shared allow long-term preservation?	A clear statement that data will be stored and shared in open formats, or in formats widely used by the community. If proprietary formats are used for data storage and sharing, information is provided justifying why open formats are not suitable and reference to software necessary to open and read these files is provided.	File formats for different data types are mentioned, but there is no indication of their suitability for long-term data preservation and sharing.	File formats and their suitability for sharing are not mentioned.
	How will you describe and document your data? Are there any metadata standards that you can adhere to in order to aid comprehension and make your data intelligible to re-users?	Clear outline of documentation and metadata strategy with references to existing good practice in the community or detailed project-specific approach where community standards do not exist.	Some mention of documentation or metadata standards without detail about community standards or a project-specific approach.	No mention of documentation or metadata.

Wellcome Rubric v2.1

When will you share your data?	A clear statement about the timeline of data sharing is made, and it is in-line with funder's policy: data will be shared no later than at the time of publication. If data sharing has to be delayed for a legitimate reason (ethical or commercial reason, or community practice), a clear justification is provided.	Timescale is mentioned but not clear or not clear for all datasets, or timescale is not in accordance with funder's expectations.	Timescale for release of data is not mentioned.
Where will you share your data?	A data repository(ies) where the data will be deposited are identified, with preference for discipline-specific repositories (if these exist). If data cannot be shared via a repository, a clear justification is provided why this is not possible.	It is indicated that a data repository will be used, but there is no mention which one, or it is explained that data cannot be shared via a repository, but there is no convincing justification why this is not possible.	The means for data sharing are not indicated.
How will you make your data accessible to others?	Clear indication how each subset of data will be available is provided, e.g. openly available, requirement for data sharing agreements, approval by data access committee.	Data re-use conditions are mentioned, but details are missing, either overall or for specific subsets of data.	Conditions for data re-use are not mentioned.
How will you make your data discoverable?	A clear plan for data discoverability is outlined, e.g. statement on data availability in publication(s), depositing data in well-indexed repository(ies) with rich metadata description, publishing a 'data paper' to accompany the dataset, creation of data catalogue record, providing reference on project/institutional websites, publicising information about the data on social media.	The intention to make data discoverable is made, but specific ways of making the data discoverable are not listed.	No mention of data discoverability.
How will you ensure that your data is properly cited?	It is indicated that persistent link(s) will be used to enable data citation or, if bespoke solutions are used, clear means of how the data will be cited are indicated.	It is mentioned that the data will get a link to enable citation, but there is no mention of persistence of that link, or if a bespoke solution for data sharing is used, no mechanism to ensure appropriate citation is indicated.	No mention of data citation.
Are there any limitations to data sharing and if so, what are they?	There is a clear assessment of any ethical, IPR or patient confidentiality concerns. Methods by which these can be mitigated are also discussed. Eg managed access to data, clear consent forms, information on NDA agreements. Alternatively, there is a clear statement that there are no restrictions on this dataset.	Data sharing restrictions or problems are mentioned and justified, but no plans to mitigate these are discussed.	Data sharing restrictions are not mentioned, or restrictions on access are indicated without justification.
How will you preserve your key datasets?	It is stated that research data will be shared via a named repository and a clear reference to the repository's preservation policy or commitment to preservation is provided. Alternatively, if bespoke solutions are used for data sharing, a reference is provided to a comprehensive policy enabling curation, preservation, long-term storage and sharing of the data.	There is no reference to the repository's preservation policy or commitment, or if a bespoke solution is used, there is no reference to policies ensuring curation, preservation, long-term storage and sharing of research data.	No re-assurance about data preservation is provided.
Will you require any resources to deliver your plan?	Required resources are listed (e.g. people infrastructure, costs of active data storage, costs of licences for software to support data management, costs of data ingestion by the repository, costs of long-term preservation) and costed (costs have been properly calculated and factored into the application), or there is a statement that no further resources are needed.	It is stated that resources are needed, but details are not provided.	No mention of required resources.