



HAL
open science

Institut Pasteur checklist for best practices in research data management

Fanny Sébire

► **To cite this version:**

Fanny Sébire. Institut Pasteur checklist for best practices in research data management. 2023. pasteur-04174938

HAL Id: pasteur-04174938

<https://pasteur.hal.science/pasteur-04174938v1>

Submitted on 1 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Institut Pasteur checklist for best practices in research data management

You can use this checklist to evaluate your practices and check that you have not forgotten anything in your data management approach.

Possible responses are the following:

Y: yes, N: no, DNK: do not know, N/A: not applicable

	Y	N	DNK	N/A
Funder requirements				
Are you aware of your research funder's requirements for data management? <i>Some funders require not only a Data Management Plan, but also to allow free reuse of data</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data Management Plan (DMP)				
Have you drawn up a DMP?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have you updated your DMP after each significant modification to the project (new data, patent application, new partners etc.)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have you appointed a person to be responsible for drawing up and updating the DMP?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Generated data				
Have you set up a single collaborative space for the storage, back-up, and sharing of datasets within your research team?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have you set up a file classification scheme for this storage space to make it easy to find your data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have you implemented procedures to ensure the security of your data (regular backup, transfer via secure tools...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do you use a naming convention to identify your data files?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
To describe your data, have you associated a metadata file with each dataset?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Y	N	DNK	N/A
Reuse of existing data				
If you are reusing existing data, do you systematically remember to identify the owner of the data and to check licenses and terms of use ? To help you, a flowchart listing the main legal issues to address before proceeding with the reuse or re-utilization of a dataset is available here .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Shareable and non-shareable data				
Have you identified which data can be freely disseminated and which data cannot be disseminated ? To help you, a flowchart listing the main legal issues to address before disseminating a dataset is available here .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
For data at risk (personal data, in particular), have you set up secure tools for data storage and exchange?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data dissemination				
Have you deposited the datasets resulting from your research in a data repository ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have you assigned a reuse license to your datasets (to protect data and tell users what they can do with it)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have you used open formats (non-proprietary and widely used) to facilitate the reuse of your data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If software is required to access and reuse your data, have you made the open source code and documentation needed to use the software available?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data archiving				
Have you identified the data to be preserved over the long term ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do you have permanent storage for your data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Source: OTELo, Inist. *Guide de bonnes pratiques : gestion et valorisation des données de recherche* (V1.1, 01/02/2016). CC-BY-NC-SA License. <https://hal.archives-ouvertes.fr/hal-01275841>



This document is made available under the terms of the [Creative Commons CC-BY-NC-SA License](https://creativecommons.org/licenses/by-nc-sa/4.0/)