

Horizon 2020 FAIR DMP Evaluation Rubric

This is an evaluation rubric for DMPs following the European Commission's [Guidelines on FAIR Data Management in Horizon 2020](#) (version 3.0. 26 July 2016). It is inspired by and based on the [example rubric](#) (CC0) from the DART project (<https://osf.io/26b9r/>), which was presented during the IDCC 2016 DMP Workshop in Amsterdam on 25 February 2016.

PERFORMANCE CRITERIA	PERFORMANCE LEVELS		
	Complete/detailed	Addressed issue, but incomplete/too vague	Did not address issue
1. States the purpose of data collection/generation.	Clearly describes why data are collected/generated and how this relates to the project's objectives.	Some information is provided about why data are collected/generated, but not enough for someone outside the project to fully understand the purpose and/or its relation to the project's objectives.	No information is provided about why data are collected/generated.
2. Describes what types of data will be collected/generated in the project.	Clearly describes what type(s) of data will be collected/generated (e.g. texts, spreadsheets, images, video/audio files, surveys, databases, field notes etc.), also distinguishing between data collection methods (experimental, observational, simulation, derived/compiled data etc.).?	Some details about data types are provided, but not enough for someone outside the project to fully understand what sort of data are involved.	No details about data types are provided.
3. Describes what file formats will be used for the data.	Clearly describes file formats in which data will be stored.	Vaguely or only partially describes file formats in which data will be stored.	Does not describe file formats used for the data.
4. Indicates whether the project will reuse existing data .	Clearly states whether any existing data will be reused, and if so, explains how.	Does not state whether any existing data will be reused. Or mentions the	There is no indication of whether any existing data are reused.

		reuse of existing data, but without explaining how.	
5. <i>Indicates the origin of the data.</i>	Clearly specifies where the data come from, how/why/by whom they were produced.	Vaguely or only partially indicates where the data come from and how/why/by whom they were produced.	Does not indicate where the data come from.
6. <i>States the expected size of the data.</i>	The expected total amount of data collected/generated (MB, GB, TB...) is clearly stated, or it is acknowledged that the data volume is not known.	The expected total amount of data is vaguely specified, and it is not clear whether the data volume is actually known. <i>E.g.: "The project will produce a large amount of data."</i>	The expected total amount of data is not specified, and there is no acknowledgement that the data volume is not known.
7. <i>Describes to whom the data might be useful.</i>	Clearly indicates by which other users or in which other contexts the data might be reused.	Vaguely states by whom or in which other contexts the data might be reused.	Does not specify by whom or in which other contexts the data might be reused.
8. <i>Indicates whether data produced or used are discoverable via metadata.</i>	Clearly indicates whether metadata are available online to aid data discovery.	Vaguely states whether online metadata are available for the data.	Does not specify whether online metadata are available for the data.
9. <i>Indicates whether data produced or used are identifiable and locatable via a standard identification mechanism.</i>	Clearly states whether a persistent and unique identifier is assigned in order to identify and locate the data, and specifies which one (e.g. a DOI).	Vaguely indicates whether a persistent and unique identifier is assigned to the data, or does not specify which one.	Does not specify whether a persistent and unique identifier is assigned.
10. <i>Describes what naming conventions are followed.</i>	Clearly describes the system or conventions used for naming data files.	Vaguely describes the file naming conventions that are followed.	Does not describe which file naming conventions are followed.

11. <i>Indicates whether search keywords will be provided to optimize reuse possibilities.</i>	Clearly states whether metadata will include search keywords to facilitate data reuse.	Vaguely states whether metadata will include search keywords.	Does not indicate whether metadata will include search keywords.
12. <i>Describes what approach to versioning will be used.</i>	Clearly describes the approach taken to easily identify different versions of the data, e.g. the use of clear version numbers.	Vaguely describes the approach used to easily identify different versions of the data.	There is no description of the approach used to identify different versions of the data.
13. <i>Specifies what metadata will be created.</i>	Clearly specifies the metadata standard(s) that will be used to describe the data. Or indicates that no metadata standards exist for the discipline, and outlines the approach used to create metadata (e.g. what type of metadata, how will they be created...).	Is vague about what metadata standard(s) will be used or whether any suitable standards exist. Or, if no suitable standards exist, only vaguely outlines the approach used to create metadata.	Does not specify what metadata standard(s) will be used, and does not outline any other approach to creating metadata.
14. <i>Identifies which data will be made openly available.</i>	Clearly states which data will be made openly accessible as the default. If (some) data cannot be shared or only under restrictions, explains why, distinguishing between legal/contractual and voluntary reasons.	Vaguely or only partially states which data will be made openly accessible, and/or does not sufficiently explain why (some) data cannot be shared or only under restrictions.	Does not indicate which data will be made openly accessible.
15. <i>Describes how data will be made available.</i>	Clearly describes how data will be made available to others, e.g. by deposition in a data repository.	Vaguely describes how data will be made available.	Does not describe how data will be made available.
16. <i>Identifies what software is needed to access the data.</i>	Clearly identifies the methods or software tools, and any documentation about the software, that are needed to access the data, and indicates whether the relevant software can be made	Vaguely or only partially identifies the software (documentation) needed to access the data, and/or does not indicate whether the relevant software can be made available along with the data.	Does not identify the software needed to access the data.

	available along with the data (e.g. as open source code).		
17. Identifies <i>where</i> research data and associated metadata, documentation, and code will be <i>deposited</i>.	Clearly states the repository where data and associated metadata, documentation and code will be deposited, clarifying whether it is certified and supports open access. Also indicates whether appropriate arrangements with the identified repository have been explored.	Vaguely or only partially states the repository where (meta)data, documentation and code will be deposited, its certification and/or support for open access. Or does not indicate whether arrangements with the identified repository have been explored.	Does not state the repository where (meta)data, documentation and code will be deposited.
18. If <i>restrictions</i> are required, specifies <i>how access</i> to the data will be provided.	Clearly describes the procedures and conditions for providing access to data that cannot be made openly available, indicating for example the use of a data access committee, clear reuse conditions (e.g. via machine-readable licenses), or any authentication procedures.	Vaguely or only partially describes the procedures and conditions for providing access to restricted data.	Does not describe how access to restricted data will be provided.
19. Indicates whether data are interoperable through the use of <i>standard (open) formats</i>.	Clearly states to what extent standard (open) formats are used that make data interoperable.	Vaguely or only partially states whether standard (open) formats are used to make data interoperable.	Does not state whether standard (open) formats are used.
20. Specifies what (meta)data <i>vocabularies, standards or methodologies</i> will be followed to facilitate interoperability.	Clearly specifies any (meta)data vocabularies, standards or methodologies that will be followed to improve data interoperability.	Is vague about (meta)data vocabularies, standards or methodologies that will be followed to improve data interoperability.	Does not specify any (meta)data vocabularies, standards or methodologies that will be followed to improve data interoperability.
21. Indicates whether standard vocabularies will be used for all data types in the dataset to allow <i>interdisciplinary interoperability</i>.	Clearly indicates whether standard vocabularies will be used for all data types in the dataset. If uncommon or project-specific ontologies or	Is vague about whether standard vocabularies will be used for all data types in the dataset, or does not indicate whether uncommon or	Does not indicate whether standard vocabularies will be used for all data types.

	vocabularies are used, indicates whether mappings to more commonly used ontologies will be used.	project-specific ontologies or vocabularies will be mapped to more commonly used ontologies.	
22. Specifies how data will be licensed for the widest possible reuse.	Clearly indicates under which license data will be made available for the widest possible reuse (e.g. a CC license).	Vaguely or only partially indicates under which license data will be made available.	Does not specify under which license data will be made available.
23. Indicates when data will be made available for reuse.	Clearly states when data will be made available for reuse. If an embargo is sought, specifies why (e.g. to give time to publish or seek patents) and how long it will apply, bearing in mind that research data should be made available as soon as possible.	Vaguely or only partially states when data will be made available for reuse. Or does not explain why an embargo period is sought and/or how long it will apply.	Does not state when data will be made available.
24. Indicates whether data produced or used in the project are usable by third parties after the end of the project.	Clearly states whether data will be usable by third parties after the end of the project. If there are reuse restrictions on some of the data, explains why.	Vaguely or only partially states whether data will be usable by third parties after the end of the project, or this can be inferred because the plan indicates under which license data will be made available for reuse. Or does not explain why reuse restrictions apply to some of the data.	Does not state whether data will be usable by third parties after the end of the project.
25. Indicates how long data are intended to remain reusable.	Clearly indicates how long data are intended to remain reusable.	Vaguely or only partially indicates how long data are intended to remain reusable, or this can be inferred because the plan indicates that data will be deposited in a repository.	Does not indicate how long data are intended to remain reusable.

26. <i>Describes the process to assure data quality.</i>	Clearly describes what processes are used to assure data quality.	Vaguely or only partially describes the processes used to assure data quality.	Does not describe what data quality assurance processes are used.
27. <i>Specifies the costs of making project data FAIR.</i>	Provides a clear estimate of the costs of making project data FAIR, and describes how these costs will be covered (e.g. costs related to Open Access to research data are eligible as part of the Horizon 2020 grant).	Is vague about the costs of making project data FAIR, and/or does not describe how these costs will be covered.	Does not provide an estimate of the costs of making project data FAIR.
28. <i>Identifies who will be responsible for data management in the project.</i>	Clearly lists the person(s) or parties responsible for each aspect of data management in the project.	Is vague about the person(s) or parties responsible for data management in the project and/or their role.	Does not list the person(s) or parties responsible for data management in the project.
29. <i>Describes the costs and potential value of long-term data preservation.</i>	Clearly specifies the costs and potential value of long-term data preservation, taking into consideration how data will be selected for preservation and for how long they will be preserved.	Vaguely or only partially specifies the costs and potential value of long-term preservation.	Does not specify the costs and potential value of long-term data preservation.
30. <i>Specifies what provisions are in place for data security.</i>	Clearly describes the provisions in place for data security, including data recovery as well as secure storage and transfer of sensitive data. Also indicates whether data are safely stored in certified repositories for long-term preservation and curation.	Vaguely or only partially describes the provisions in place for data security. And/or does not indicate whether data are safely stored in certified repositories for long-term preservation and curation, or this can be inferred because the plan indicates deposit in a data repository.	Does not describe the provisions in place for data security, nor indicate whether data are safely stored in certified repositories.
31. <i>Indicates whether there are any ethical or legal issues that have an impact on data sharing.</i>	Clearly states whether any ethical or legal issues have an impact on data sharing, or indicates that this is discussed in the context of the ethics review. If relevant, includes references	Vaguely or only partially states whether any ethical or legal issues have an impact on data sharing, or whether this is discussed in the ethics review.	Does not state whether any ethical or legal issues have an impact on data sharing, nor include references to the ethics review,

	to ethics deliverables and the ethics chapter in the Description of the Action (DoA).		ethics deliverables or the ethics chapter in the DoA.
32. <i>If dealing with personal data, indicates whether data sharing and long-term preservation is included in informed consent forms.</i>	Clearly states whether informed consent for data sharing and long-term preservation is sought.	Vaguely states whether informed consent for data sharing and long-term preservation is sought.	Does not state whether informed consent for data sharing and long-term preservation is sought.
33. <i>Indicates whether other national/funder/sectorial/departmental procedures for data management are used.</i>	Clearly indicate whether other national, funder, sectorial or departmental procedures for data management are followed and if so, specifies which ones.	Vaguely indicates whether other national, funder, sectorial or departmental procedures are followed, or does not specify which ones.	Does not indicate whether other procedures for data management are followed.