

Data Set	DATA MANAGEMENT PLAN FOR H2020 PROJECTS				
No.	Partner Legal Name	Project/GA No/WP	Researcher's name	email	
Please choose from available lists or add options if applicable.				NOTES	
1. DATA SUMMARY					
	Purpose of the Data	Purpose	Objectives	State the purpose of the data collection/generation, indicating the relation with the objectives of the project. Add additional objectives if necessary.	
			1		
			2		
			3		
	Type and Format of Data	Form	Format	Describe the type of data used or generated within the project, specifying the form and format of the data.	
	Text	type_of_text	type of format	Form: Field or laboratory notes, survey responses .. Format: in plain text, (txt), HTML, XLM, PDF/A ...	
		type_of_text	type of format		
		type_of_text	type of format		
	Numeric	tables		Tables, row counts, measurements - in .XLSX, .CSV ...	
		tables			
	Audiovisual	image	jpg	Images, sound recordings, video - in .JPEG, .JPG, .PNG, .TIFF, AIFF, WAVE, .MP3, .MP4...	
		image	jpg		
	Simulated				Please state the model, model type and computer code - and specify output data type and format.
	model				
	model type				
	computer code				
	data type				
	format				
	Discipline specific information	discipline	format	e.g.: CIF in chemistry ... (specify discipline and format)	
economics					
Instrument specific	equipment	format	Equipment output (specify equipment and format).		
Reused-Data (rd)	yes_rd	write explanation	Indicate if you re-use existing data (generated outside the Oyster project). If so, explain how.		
	Data Origin	Define and describe the origin/source of your data. Data can be gathered from different sources.			
	Observational			Data captured in real time - often not reproducible i.e. sensor readings, images, telemetries, sample data...	
	Experimental			Data from lab equipment, often reproducible, but with high costs - i.e. chromatograms, magnetic fields readings...	

	Simulation			<i>Data generated by computational models where model and metadata are equally important to output data - i.e. climate models, economic models, materials models,...</i>
	Derived/Compiled			<i>Data coming from analysis or compilation. Reproducible but with high costs - i.e. the results of text and data mining, compiled databases..</i>
	Reference or Canonical (links)		<i>write the reference or canonical</i>	<i>Collection or conglomeration of smaller (peer-reviewed) datasets published and curated - i.e. chemical structures, gene sequence databanks, spatial data portals..</i>
	Dataset is:			Fixed : never change after being collected or generated. Growing : new data may be added, but the old data is never changed or deleted. Revisable : new data may be added, and old data may be changed or deleted.
	Quantity		<i>in MB/GB</i>	<i>of each experiment</i>
			<i>in MB/GB</i>	<i>overall</i>
				<i>In case not just digital archiving is required, indicated quantities of other form of storage.</i>
	Data Security & Storage	<i>select or add: type of storage</i>	<i>write data security policy</i>	<i>(i.e. Office computer, Hard Drive, Tape back-up system, Institute network drive, Institute Central Data storage, private Cloud storage ...), briefly describing the data security policy applied.</i>
	Data Value (long term)			<i>Describe to whom the data could be useful.</i>
				<i>Estimate potential value of long-term re-use of the data.</i>
2 FAIR DATA				
2.1 FAIR DATA - Making data findable				
	Discoverability of data (metadata provision)	<i>write explanation</i>	<i>write information</i>	<i>Explain how data are documented and if metadata are provided, listing the information made available/discoverable.</i>
	Identifiability of data (refer to standard id mechanisms)	<i>choose mechanism</i>	<i>write how are made identifiable</i>	<i>Indicate how data are made identifiable, if a standard permanent identifier assignation scheme is used (i.e. ARK, DOI, PURL, URN, MODA ...)</i>
	Naming conventions used	<i>describe</i>	<i>refer</i>	<i>Describe the system used to name and structure electronic files and folders. Refer also to any file renaming procedure or tools used.</i>
	Search keywords approach	<i>indicate</i>		<i>Indicate the approach to keywords generation, indexing and tagging. (For materials modelling the MODA provide this answer.)</i>

	Clear versioning approach	Versioning	Traceability	Describe the versioning and traceability approach used (especially if the dataset is growing or revisable).
		<i>indicate</i>	<i>describe</i>	Indicate and describe the procedures and templates applied for the creation of metadata.
	<i>refer</i>	<i>initiatives</i>	Refer to any institute policy or recommendations by specific initiatives that are applied.	
	Standards or procedures for metadata creation applied			Some references: MODA, EMMO (European Materials Modelling Ontology), Dublin Core Metadata Initiative, DataCite Metadata Schema, Open Archives Initiative Object Reuse and Exchange, ISAtools ... If there are no standards in your discipline, describe what type of metadata will be created and how.

2.2 Fair data Making data openly accessible

	Data openly available	<i>indicate ownership</i>		Indicate ownership of the data, if it is openly available or can be made openly available.
	Data kept closed	<i>users</i>	<i>reasons</i>	Indicate if data access is restricted, to what users, and explain the reasons.
	How data will be made available	<i>indicate</i>		Indicate how you intend to make data available.
	Methods or software (SW) tools for data access	<i>write methods and tools</i>		Indicate methods and SW tools needed to access the data. Clarify if the relevant software (e.g. in open source code) is included in the data set.
	SW documentation and other information needed	<i>indicate</i>		Indicate any specific SW documentation that is needed to access the data, or additional information that is needed to understand the data (i.e. abbreviations, supplementary notes).
	Repository for deposit of data, metadata, documentation and code	<i>indicate open or private</i>		Indicate the (open or private) repositories in which the data, metadata, documentation and code are stored and/or those in which they will be stored in the future.
	Access restrictions	<i>indicate</i>	<i>explain</i>	Indicate if there are limitations and restrictions to access the data, and if they are linked to a specific timeframe. Explain how access will be provided after these restrictions are lifted.

	Data interoperability assessment	<i>indicate</i>	<i>range of utilization</i>	<p>Assess the level of interoperability of the dataset.</p> <p>Indicate data and metadata vocabularies, standards and methodologies followed to facilitate interoperability.</p> <p>Indicate if open standards are used, and (if you know) the range of utilization of proprietary SW and methodologies used to generate and manage the data.</p>
		<i>select</i>		
2.3 Fair data Making data interoperable				
	Standard vocabulary or mapping to commonly used ontologies	<i>refer</i>		<p>Refer to commonly used ontologies to map the dataset, considering also the use of existing common platforms and tools – e.g.: EMMO, BFO, MatONTO, Materials Ontology..</p>
	Data licensing for wide reuse	<i>define</i>	<i>indicate</i>	<p>If applicable, define data licensing approach for the dataset wide reuse. Indicate the chosen licenses tools.</p>
2.4 FAIR DATA – Increase data re-use (through clarifying licenses)				
	Timing of data availability for re-use (incl. indications on embargo)	<i>define</i>	<i>indicate</i>	<p>If applicable, define the timeframe for making data available for re-use. Indicate any embargo period if required.</p>
	Data usability by Third Parties (after the end of the project)	<i>indicate</i>		<p>Indicate any limitation to the use of the data by Third Parties, after the end of the project.</p>
	Restrictions to data re-use	<i>indicate</i>	<i>explain</i>	<p>Indicate and explain any restriction to the re-use of data (i.e. confidentiality agreements, other issues).</p>
	Quality assurance process	<i>how assured</i>	<i>how controlled & documented</i>	<p>Explain how quality of the data is assured, how the consistency and quality of data collection is controlled and documented.</p>
	Length of time of data re-usability	<i>indicate</i>		<p>Indicate the time limit for the data re-usability, if any.</p>
3 ALLOCATION OF RESOURCES				
	Costs estimates for making data FAIR	<i>estimate</i>	<i>describe</i>	<p>Estimate the costs for making your data FAIR (findable, accessible, interoperable and reusable) and describe how you intend to cover these costs (i.e. institute dedicated resources, dedicated part of the project budget ...).</p>
	Data Management Responsibilities	<i>identify</i>		<p>Identify responsibilities for data management of this dataset (within your research group and institute, and within the project if applicable).</p>