COASTAL
Collaborative Land-Sea Integration Platform

D26 Information & Data Management Plan

Version Final
WP 7, T 7.3 Lead beneficiary: VITO

Authors: Jean-Luc de Kok (VITO)
Aris Karageorgis (HCMR)
Samaneh Seifollahi (SU)
Rachel Tiller (SINTEF)
Ebun Akinsete (ICRE8)
Katalin Balazs (GEO)
Noémie Wouters (GRBR)

Approved by WP Manager:	Jean-Luc de Kok (VITO)
Date of approval:	October, 2018
Approved by Project Coordinator:	Jean- Luc De Kok(VITO)
Date of approval:	October, 2018
Due date of deliverable:	October 31, 2018
Actual submission date:	October 29, 2018

H2020-RUR-02-2017
Project number: 773782
Project duration: 1 May 2018 – 30 April 2022
Project coordinator: Jean Luc De Kok (VITO)

Project web site: www.h2020-coastal.eu

COASTAL Knowledge Exchange Platform: www.coastal-xchange.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773901.

Dissemination level		
PU	Public	**X**
CO	Confidential, restricted under conditions set out in Model Grant Agreement	
CI	Classified information as referred to in Commission Decision 2001/844/EC)	
Deliverable type		
R	Document, report	**X**
DEM	Demonstrator, pilot, prototype	
DEC	Web sites, patent fillings, videos, etc.	
OTHER	Software, technical diagram, etc.	

History			
Version	Date	Reason	Revised by
0.1	14.08.18	First draft	Jean-Luc de Kok
0.2	16.08.18	Second draft	Jean-Luc de Kok
0.3	20.08.18	Third draft	Jean-Luc de Kok
0.4	21.08.18	OpenAire clarification included	Jean-Luc de Kok
0.5	31.08.18	General update	Jean-Luc de Kok
1.0	16.10.18	Final update	Jean-Luc de Kok
1.3	19.10.18	Comments and corrections HCMR and SU processed.	Aris Karageorgis, Samaneh Seifollahi, Jean-Luc de Kok
2.0	22.10.18		Jean-Luc de Kok
Final	29.10.2018	Final version completed	Jean-Luc de Kok

Partnership:

This document was produced under the terms and conditions of Grant Agreement No. 773782 for the European Commission. It does not necessary reflect the view of the European Union and in no way anticipates the Commission's future policy in this area

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

Executive summary

In H2020 projects a Data Management Plan (DMP) describes the data management life cycle for all datasets collected, including the handling of research data during and after the project. This deliverable is a draft DMP, which will be updated during the project lifetime when significant changes demand this. Here Open Data refers to publicly available, non-personal, scientific data which can be machine read, comply to agreed standards and can be (re)used without costs if properly cited, for commercial and non-commercial use. The Open Research Data Pilot (ORDP) is a natural framework for dissemination and exploitation of the main project outcomes, aiming to improve and maximize access to and reuse of research data generated by the project. Efficient and proper data management is an important aspect of COASTAL, reflected by the project title ('knowledge exchange') and the consortium agreed to participate in the Open Research Data Pilot of Horizon 2020 and FAIR principles for use of metadata[1]. Research data collected by WP2 (Knowledge Transition) to support the systems modelling and policy/business analyses will be harmonised. These data include references to existing environmental and social-economic research data and also newly generated data by the partners of COASTAL. The project information management (Task 7.3) is centralised with the Project Coordinator, with the input from WP leaders and local research partners. Each partner is held responsible for keeping proper administrative information and financial data about the project. The preparation of a proper DMP is essential within the context of the broader task of Project Information Management. The DMP for COASTAL follows the template included in the H2020 Online Manual and is fixed into this Final Data Management & Information Plan. Each WP lead partner and local R&D partner is asked to keep records of information and data generated during the project implementation. The PMT (WP lead partners and coordinator) has been tasked with the responsibility of labelling each item generated according to its nature (e.g. strictly confidential as opposed to public). The PMT will also decide on which information will be made publicly available on the website. A specific session within the PMT meetings will be dedicated to Information management. The handling of personal data, which is relevant in the context of the new EU GDPR Regulation and the activities in the COASTAL case studies, is addressed in detail in Ethics deliverable D1 – Ethics Requirements No. 1 but summarised in this document.

[1] https://www.force11.org/group/fairgroup/fairprinciples

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

List of Tables

List of Figures

Abbreviations and acronyms

CA	Consortium Agreement
CMS	Content Management System
DMP	Data Management Plan
DPO	Data Protection Officer
FAIR	Findable, Accessible, Interoperable, Reusable
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
MAL	Multi-Actor Lab
OpenAIRE	Open Access Infrastructure for Research in Europe
ORDP	Open Research Data Pilot

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

1. Introduction

Under the H2020 Programme, Open Access[2] refers to the on-line access to scientific information, free of charge and in a reusable format, as specified in Articles 29.2 and 29.3 of the Annotated Grant Agreement[3]. This includes scientific publications as well as **research data**. The **Data Management Plan** (DMP) handles the life cycle of all research data with **potential for reuse,** collected during and after the project. Open access to such data refers to publicly available, non-personal, data which can be machine read, comply to agreed standards and can be (re)used without costs if properly cited, for commercial and non-commercial use, and following certain requirements for use . See, for example, https://statbel.fgov.be/en , an open databank for statistics for Belgium. There are several reasons why Open Access is a good practice and of growing interest:

- Fostering scientific progress and making science more efficient (not 'reinventing the wheel')
- Economic growth resulting from the use of well-managed, publicly available, data
- A broad, fast and more transparent access to data by researchers, citizens and industry
- A benefit to the career of the data developers including, for example, citations

A general misconception among researchers is that Open Access affects the **protection of knowledge** resulting from their research efforts, limiting the exploitation of their work. This is incorrect:

- The principle is "as open as possible, as closed as necessary".
- There is no (formal) requirement to publish under H2020, however, if research is published this should be done as open access publications.
- Research can still be exploited commercially, e.g. through patenting.
- It is possible to postpone publishing, and first seek protection.
- Background data (i.e. data developed outside the scope of the project or before the project) can be protected automatically as part of the Consortium Agreement.
- Finally, users are generally more interested in the use of the data and related expertise rather than the data themselves – service oriented business models, as foreseen for COASTAL, can meet this demand and still ensure benefits to the data developers and their knowledge (see Figure 1).

The Open Research Data Pilot (ORDP) aims to make the research data generated by Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access. The ORDP has been extended to all thematic areas and covers:

[2] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

[3] http://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

- Underlying data needed to validate results of scientific publications
- All other data, as specified in the DMP, which focuses on research data with potential for reuse

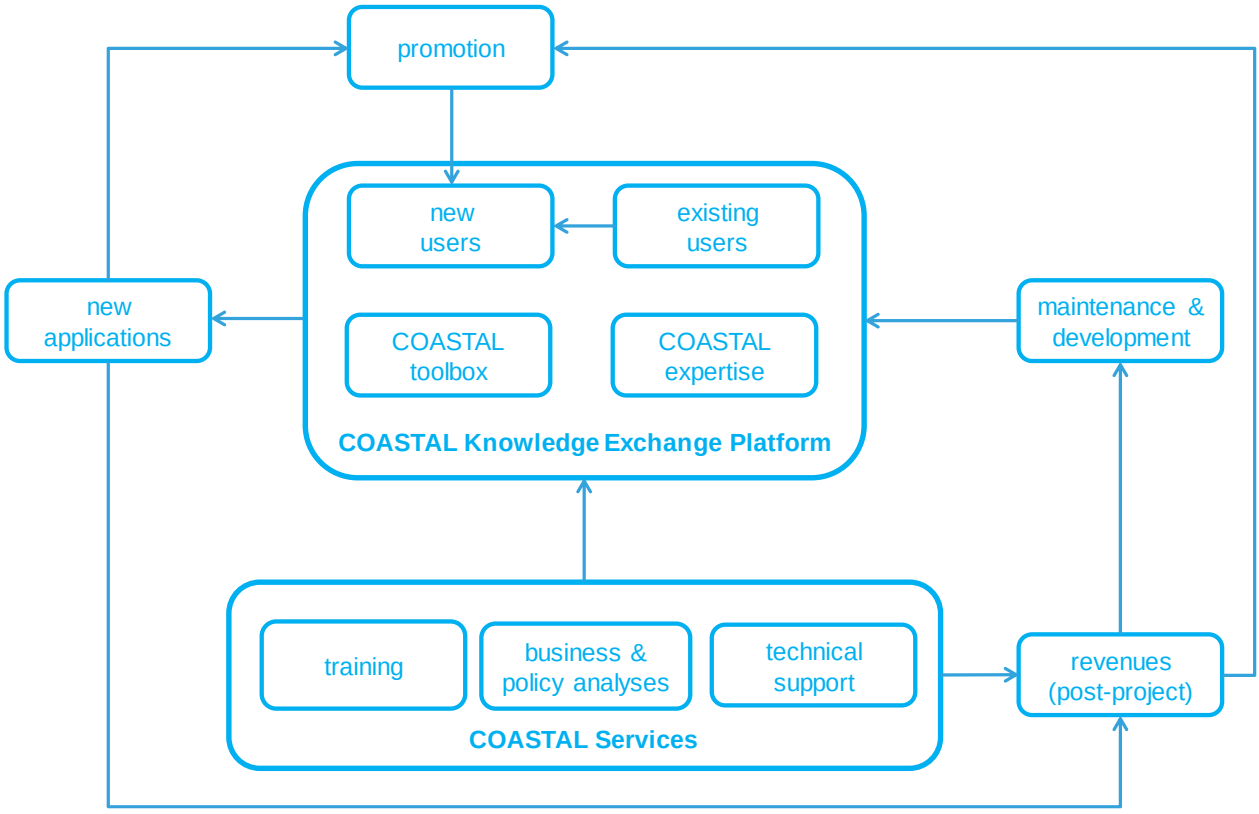

Figure 1 Service-oriented business model foreseen for COASTAL after the project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

2 Data Management and the FAIR Principles

The purpose of the DMP is to anticipate on problems or challenges encountered with data handling and will help identify ways to overcome problem. It is a living document – this deliverable includes the first draft. The DMP should include[4]:

- A description of the data: what kind of data will the project collect or generate, and to whom might they be useful later on? The Open Research Data Pilot (ORDP) applies to (1) the data and metadata needed to validate results in scientific publications and (2) other curated and/or raw data and metadata that may be required for validation purposes or with reuse value.
- Standards and metadata: relevant disciplinary standards adopted by the project; data context; data creators and purpose of developing the data; data format. Metadata, documentation and standards should be included to make the data Findable, Accessible, Interoperable and Re-usable (FAIR).
- Data sharing: as much of the resulting data as possible should be archived as Open Access, and reasons for not sharing resulting data should be explained. Intellectual Property Rights (IPR) agreements, legal and ethical factors impacting data sharing can be taken into consideration and should be explained.
- Archiving and preservation: publicly funded research output should have a positive impact on future research, for policy development, and for societal change. Data are to be made available for a suitable period beyond the project lifetime. Data storage should take into consideration code (software) used, and other tools needed to validate and use the preserved data.

Practically, three steps are to be taken[5]:

1. The creation of the DMP (see Section 4)
2. Identifying an open data repository to preserve the data, metadata and tools in the long term. In the case of COASTAL a project-related Knowledge Exchange Platform will be set up with a similar purpose (i.e. providing free access to project results and tools in a structured way). Therefore, the PMT (WP lead partners and coordinator) should carefully examine the different options against one another.
3. The data selected for open access should be selected based on their usefulness for the project consortium and other parties, to avoid spending effort on data which are of little use.

To ensure the data sharing is FAIR[6] (Findable, Accessible, Interoperable and Reusable) it is recommended to:

[4] https://www.openaire.eu/what-is-the-open-research-data-pilot

[5] https://www.openaire.eu/what-is-the-open-research-data-pilot

[6] https://www.force11.org/group/fairgroup/fairprinciples

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

- Refer to appropriate data repositories (www.re3data.org)
- Document and describe the data to allow understanding, reuse and citing by others (www.dcc.ac.uk/resources/how-guides/cite-datasets)
- License the data for reuse by others (www.dcc.ac.uk/resources/how-guides/license-research-data)

Participating in the ORDP does not necessarily mean opening up all research data. Rather, the ORDP follows the principle "**as open as possible, as closed as necessary**" and focuses on encouraging sound data management as an essential part of research best practice. Where data need to be shared under restrictions, explain why, clearly separating legal and contractual reasons from voluntary restrictions. It is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

3 Data Management in COASTAL

The COASTAL consortium agreed to participate in the ORDP of Horizon 2020 and FAIR principles for access to data. The general project framework is based on a combination of systems thinking/modelling and multi-actor analysis. It is anticipated that a proportion of the supportive data for the modelling will only be made available to the project under certain restrictions. When these apply to the use of existing field data, statistics, maps etc. these data can be protected by specific conditions of use for the COASTAL Platform, such as proper citing of the data owners. Actor partners and stakeholders engaged in the Multi-Actor Labs will be granted access to, models and scenarios available on the knowledge exchange platform (https://coastal-xchange.eu) depending on their needs and the state of completion of these data. For this purpose, the COASTAL web site (https://h2020-coastal.eu) and knowledge exchange platform will come with functionalities allowing for restricted access for a group of users involved in a project. The type of data, data ownership, management, use, distribution and access rights to the data are to be described for every data provider in the DMP.

The objective of WP2 – Knowledge Transition - is to develop the quantitative data and scientific model constructs needed for analyzing the rural-coastal interactions identified in WP1 – Multi-Actor Analysis . These need to be translated to the appropriate level of detail required for strategic business and policy analysis. To the extent possible, the focus of WP2 will be on the translation of existing data and models for the quantification of the social-economic, physical, and environmental interactions in a System Dynamics (SD) model as input to WP4 – Systems Modelling. WP2 will connect with WP1 and WP4 to delineate the problem context, potential causes, and spatial-temporal boundaries and resolution for the modelling based on the outcomes of the sectoral workshops with local actors and stakeholders. The outcomes are used to identify the types of data and the associated models required for analysing the coastal and rural processes for each case study. A **Data and Model Inventory** (deliverable D6, due M6) will be prepared for each Multi-Actor Lab (MAL). This inventory will include guidelines on how to address the modelling of interactions for which models and data are not available.

Together, the Data and Model Inventory and this **DMP** will specify:

the type of data collected and/or generated by the project (see

- Table 1) that will be made public;
- the procedures planned for making these data accessible for verification and reuse;
- the procedures planned for making other data restricted under some conditions;
- the procedures planned for curing and preserving data.

The available information will be managed and stored in a Content Management System (CMS – provided and maintained by GEO), taking advantage of existing information management open sources that could be adaptable to project data dissemination needs. It will include both, Publications and Repository of other

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

research data. Open access to research data refers to right to access and reuse digital research data under the terms and conditions set out in the Grant Agreement. Openly accessible research data can typically be accessed, mined, exploited, reproduced, and disseminated free of charge for the user.

Table 1 COASTAL data typology.

WP	main data delivered	preferred format	developing partners	main users
1	mental maps; fuzzy cognitive maps	*.cmap; *.vmf; *.xlsx; *.m (MatLab); R	4	WP2, WP4
2	model metadata; source data/model inventory	*.xlsx; report (D2.1)	1-3, 5-9, 11 + actors	WP4
3	business road maps; policy guidelines; indicator time series	report (D3.3); report (D3.3); *.xlsx	all	WP1, WP6
4	SD (sub)models; generic response functions; parameter settings	*.stx or *.mdl; *.xlsx; *.txt	1-3, 5-9	WP3, WP6
5	scenarios; transition pathways	*.xlsx; report (D5.1)	8, 11	WP3, WP4
6	webpages, templates	*.docx; *.xlsx; *.html	1,8, 10	All partners

The Consortium Agreement (CA) for COASTAL distinguishes project foreground and backgrounds:

- Foreground: models and data developed as part of the project, automatically subject to the principle of Open Access
- Background: all other models and data developed and/or owned by the project beneficiaries. Under certain circumstances it can be useful to provide access to background material to other partners or third parties. In the CA the project partners identified and agreed on the Background for the Project and have also, where relevant, informed each other that Access to specific Background is subject to legal restrictions or limits. Anything not identified as Background is not the object of Access Right obligations regarding Background. These access rights can be modified in case it is desirable to open or close access the Background.

The DMP should identify results subject of dissemination and exploitation activities. It will also analyse the main data uses and users and explore the restrictions related to IPR according with the CA.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

4. The Data Management Plan (DMP)

A template will be used to support the implementation of the DMP at the level of the individual MALs (see Table 2). It is based on the template available for the H2020 Programme[7].

Table 2 Template for the data management.

Data Management Plan (DMP)	Issues addressed
1. Data summary	• State the purpose of the data collection/generation • Explain the relation to the objectives of the project • Specify the types and formats of data collected/generated • Specify if existing data is being reused (if any) • Specify the origin of the data • State the expected size of the data (if known) • Outline the data utility: to whom will it be useful
2. FAIR Data 2.1. Making data findable, including provisions for metadata	• Outline the discoverability of data (metadata provision) • Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? • Outline naming conventions used • Outline the approach towards search keyword • Outline the approach for clear versioning • Specify standards for metadata creation (if any). If there are no standards in your discipline, describe what type of metadata will be created and how
2.2. Making data openly accessible	• Specify which data will be made openly available? If some data is kept closed, provide rationale for doing so • Specify how the data will be made available • Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)? • Specify where the data and associated metadata, documentation and code are deposited

[7] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

	• Specify how access will be provided in case there are any restrictions
2.3. Making data interoperable	• Assess the interoperability of the data. Specify what data and metadata vocabularies, standards or methodologies will be followed to facilitate interoperability. • Specify whether standard vocabulary will be used for all data types present in the data set, to allow inter-disciplinary interoperability? If not, will mapping be provided to more commonly used ontologies?
2.4. Increase data reuse (through clarifying licenses)	• Specify how the data will be licensed to permit the widest reuse possible • Specify when the data will be made available for reuse. If applicable, specify why and for what period a data embargo is needed • Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the reuse of some data is restricted, explain why • Describe data quality assurance processes • Specify the length of time for which the data will remain reusable
3. Allocation of resources	• Estimate the costs for making your data FAIR. Describe how these costs will be covered • Clearly identify responsibilities for data management in the project • Describe costs and potential value of long term preservation
4. Data security	• Address data recovery as well as secure storage and transfer of sensitive data
5. Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former
6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that are used (if any)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

5 Handling of personal and sensitive data

The new EU Regulation 2016/679[8] builds on the earlier General Data Protection Regulation (95/46/EC) or GDPR and is aimed at ensuring the protection of natural persons with regard to the processing of personal data. **Personal data** refers to information corresponding to a natural person (a so-called 'data subject') who has been or can be identified directly, or indirectly. In particular this will be the case if identifiers such as names, addresses, id numbers etc. are used. It also refers to person-specific factors such as gender, physical, mental, social-economic and cultural characteristics. The **data processing** refers to the collection, recording, storage, adaptation, disclosure, forwarding, destruction, and all uses of the data.

The **key principles** are:

- The personal data should be adequate, relevant and limited for the intended use, rather than what data are desirable to collect and process
- Accountability of the data processing and data replaces a compliance-based approach (record keeping)
- It is mandatory to appoint an independent data protection officer, who will provide advice and can evaluate the data processing against compliance with the GDPR.
- The rights of the data subjects are strengthened, for example natural persons should give explicit consent before personal data can be collected, and the data can be removed upon request.

The general approach in COASTAL of handling personal data is based on two principles:

- **Anonymization**: this means that personal data are processed in a way that makes the risk of identification negligible. For example, interview data can be collected without storing the names, gender, age of the data subjects responding to the survey.
- **Pseudo-anonymization**: here direct identifiers such as names, addresses, affiliations, etc. are replaced by indirect identifiers (for example numbers) in the data set. A separate data set, known as the 'key', is used to link the indirect identifiers to the direct identifiers.

More specific measures taken in COASTAL to ensure compliance with the GDPR are:

- to ensure all project partners will be **informed** and **reminded regularly** of the obligations with respect to the collection and processing of personal data related to EU Regulation 2016/679
- implement the **technical and organisational measures** to ensure the collection and processing of personal data are transparent, lawful, and limited to the purposes specified to the data subjects (natural persons) following (EU Reg. 2016-679 – Art. 28)

[8] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

- a central, independent **Data Protection Officer** (DPO) has been appointed for the duration of the project to monitor compliance with the Regulation and provide advice to the project consortium members (EU Reg. 2016-679 – Art. 35)
- ensure the agreement of data subjects to the storage and processing of personal data will be based on **clear, affirmative consent** through verifiable electronic or written means (EU Reg. 2016-679 – Art. 7)

Personal data will be collected primarily by Partner 4 (SINTEF) and Partner 10 (GEO), related to the activities in WP1 (Multi-Actor Analysis) and WP6 (COASTAL web site), and other partners to support the administration of events with stakeholders and actors. An overview of the types of personal data to be collected by the beneficiaries concerned (partners 1, 4, 8, and 10 in particular) is provided in **Error! Reference source not found.** of Deliverable D1 – Ethics Requirements.

The handling of sensitive research data has been foreseen in the DMP (Table 2) which will be provided to the Multi-Actor Labs and is to follow the guidelines provided in this document (notably the Introduction Section) and project CA. This type of data will primarily consist of raw research data, third party data and publishable but unscreened project outcomes (foreground). The design of the COASTAL Knowledge Exchange Platform (https://coastal-xchange.eu/) will have a key role in this respect and will take into consideration the principles outlined in this document.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782

6 Useful web links

- EC's Guide on Open Access to Scientific Publications and Research Data in Horizon 2020 (updated August 25, 2016)
  http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- EC's Guidelines on Data Management in Horizon 2020 (updated July 26, 2016):
  http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
  EC's Agenda on Open Science: https://ec.europa.eu/digital-agenda/en/open-science
- DMPonline tool: https://dmponline.dcc.ac.uk/
- DCC How to Write a DMP guide: http://www.dcc.ac.uk/resources/how-guides/develop-data-plan
- DCC How to Select What Data to Keep guide: http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep
- DCC How to License Research Data guide: http://www.dcc.ac.uk/resources/how-guides/license-research-data
- RDNL video The what, why and how of data management planning:
  http://datasupport.researchdata.nl/en/start-de-cursus/ii-planfase/datamanagementplanning/

  Software Sustainability Institute's Software Management Plan:
  https://www.software.ac.uk/sites/default/files/images/content/SMP_Checklist_2016_v0.1.pdf
- AddThis Sharing Buttons
- Share to TwitterShare to FacebookShare to LinkedInShare to Google+Share to PinterestShare to EvernoteShare to OneNoteShare to EmailShare to PrintShare to More
- Catalogue of data repositories: www.re3data.org
- Registry of Open Access Repositories: http://roar.eprints.org
- Directory of Open Access Repositories: http://www.opendoar.org
- FOSTER project on resources on Opn Science: www.fosteropenscience.eu
- Citing open data in a transparent way: www.dcc.ac.uk/resources/how-guides/cite-datasets
- Licensing your data to allow reuse by 3rd parties: www.dcc.ac.uk/resources/how-guides/license-research-data
- License Wizzard: https://eudat.eu/
- Metadata standards: http://rd-alliance.github.io/metadata-directory/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773782