Ref. Ares(2018)3869358 - 20/07/2018
Co-funded by the Horizon 2020 programme of the European Union

DOCKSTHEFUTURE
defining the concept of "Port of the Future"

Data Management Plan

Deliverable 6.6

Date: 16th July 2018

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 770064

Document Status	
Deliverable Lead	Circle
Internal Reviewer 1	n.a.
Type	Deliverable
Work Package	WP6: Project Management
ID	Data Management Plan
Due Date	1st July 2018
Delivery Date	19th July 2018
Status	Final version
Dissemination Level	Public

Document History	
Contributions	Circle
Final Version	Circle

Disclaimer

The views represented in this document only reflect the views of the authors and not the views of Innovation & Networks Executive Agency (INEA) and the European Commission. INEA and the European Commission are not liable for any use that may be made of the information contained in this document. Furthermore, the information are provided "as is" and no guarantee or warranty is given that the information fit for any particular purpose. The user of the information uses it as its sole risk and liability

Table of Contents

1. Executive summary

The deliverable outlines how the data collected or generated will be handled during and after the DocksTheFuture project, describes which standards and methodology for data collection and generation will be followed, and whether and how data will be shared.

The purpose of the Data Management Plan (DMP) is to provide an analysis of the main elements of the data management policy that will be used by the Consortium with regard to the project research data. The DMP covers the complete research data life cycle. It describes the types of research data that will be generated or collected during the project, the standards that will be used, how the research data will be preserved and what parts of the datasets will be shared for verification or reuse. It also reflects the current state of the Consortium Agreements on data management and must be consistent with exploitation.

This Data Management Plans sets the initial guidelines for how data will be generated in a standardised manner, and how data and associated metadata will be made accessible. This Data Management Plan is a living document and will be updated through the lifecycle of the project.

2. EU LEGAL FRAMEWORK FOR PRIVACY, DATA PROTECTION AND SECURITY

Privacy is enabled by protection of personal data. Under the European Union law, personal data is defined as "any information relating to an identified or identifiable natural person". The collection, use and disclosure of personal data at a European level are regulated by the following directives and regulation:

- Directive 95/46/EC on protection of personal data (Data Protection Directive)
- Directive 2002/58/EC on privacy and electronic communications (e-Privacy Directive)
- Directive 2009/136/EC (Cookie Directive)
- Regulation 2016/679/EC (repealing Directive 95/46/EC)
- Directive 2016/680/EC

according to the Regulation 2016/679/EC, **personal data**

means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person (art. 4.1).

The same Directive also defines personal **data processing** as

any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction (art. 4.2).

3. Purpose of data collection in DocksTheFuture

This Data Management Plan (DMP) has been prepared by taking into account the template of the "Guidelines on Fair Data Management in Horizon 2020" (http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf). According to the latest Guidelines on FAIR Data Management in Horizon 2020 released by the EC Directorate-General for Research & Innovation "beneficiaries must make their research data findable, accessible, interoperable and reusable (FAIR) ensuring it is soundly managed".

The elaboration of the DMP will allow to DTF partners to address all issues related with ethics and data. The consortium will comply with the requirements of Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

DocksTheFuture will provide access to the facts and knowledge gleaned from the project's activities over a two-year and a half period and after its end, to enable the project's stakeholder groups, including creative and technology innovators, researchers and the public at large to find/re-use its data, and to find and check research results.

The project's activities aim to generate knowledge, methodologies and processes through fostering cross-disciplinary, cross-sectoral collaboration, discussion in the port and maritime sector. The data from these activities will be mainly shared through the project website. Meeting with experts and the main port stakeholders will be organised in order to get feedback on the project and to share its results and outcomes.

DocksTheFuture will encourage all parties to contribute their knowledge openly, to use and to share the project's learning outcomes, and to help increase awareness and adoption of ethics and port sustainability.

4. Data collection and creation

Data types may take the form of lists (of organisations, events, activities, etc.), reports, papers, interviews, expert and organisational contact details, field notes, quantitative and qualitative databases, videos, audio and presentations. Video and Presentations dissemination material will be made accessible online via the DocksTheFuture official website and disseminated through the project's media channels (Twitter, LinkedIn and Facebook), EC associated activities, press, conferences and presentations.

DocksTheFuture will endeavour to make its research data 'Findable, Accessible, Interoperable and Reusable (F.A.I.R)', leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.

The DocksTheFuture consortium is aware of the mandate for open access of publications in the H2020 projects and participation of the project in the Open Research Data Pilot.

More specifically, with respect to face-to-face research activities, the following data will be made publicly available:

- Data from questionnaires in aggregate form;
- Visual capturing/reproduction (e.g., photographs) of the artefacts that the participants will co-produce during workshops.

5. Data Management and the GDPR

In May 2018, the new European Regulation on Privacy, the General Data Protection Regulation, (GDPR) came into effect. In this DMP we describe the measures to protect the privacy of all subjects in the light of the GDPR. All partners within the consortium will have to follow the same new rules and principles.

In this chapter we will describe how the founding principles of the GDPR will be followed in the Docks The Future project.

Lawfulness, fairness and transparency

Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject.

All data gathering from individuals will require informed consent individuals who are engaged in the project. Informed consent requests will consist of an information letter and a consent form. This will state the specific causes for the activity, how the data will be handled, safely stored, and shared. The request will also inform individuals of their rights to have data updated or removed, and the project's policies on how these rights are managed. We will try to anonymise the personal data as far as possible, however we foresee this won't be possible for all instances. Therefore further consent will be asked to use the data for open research purposes, this includes presentations at conferences, publications in journals as well as depositing a data set in an open repository at the end of the project. The consortium tries to be as transparent as possible in their collection of personal data. This means when collecting the data information leaflet and consent form will describe the kind of information, the manner in which it will be collected and processed, if, how, and for which purpose it will be disseminated and if and how it will be made open access. Furthermore, the subjects will have the possibility to request what kind of information has been stored about them and they can request up to a reasonable limit to be removed from the results.

Purpose limitation

Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes.

Docks The Future project won't collect any data that is outside the scope of the project. Each partner will only collect data necessary within their specific work package.

Data minimisation

Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.

Only data that is relevant for the project's questions and purposes will be collected. However since the involved stakeholders are free in their answers, this could result in them sharing personal information that has not been asked for by the project. This is normal in any project relationship and we therefore chose not to limit the stakeholders in their answer possibilities. These data will be treated according to all guidelines on personal data and won't be shared without anonymization or explicit consent of the stakeholder.

Accuracy

Personal data shall be accurate and, where necessary, kept up to date

All data collected will be checked for consistency.

Storage limitation

Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed

All personal data that will no longer be used for research purposes will be deleted as soon as possible. All personal data will be made anonymous as soon as possible. At the end of the project, if the data has been anonymised, the data set will be stored in an open repository. If data cannot be made anonymous, it will be pseudonymised as much as possible and stored for a maximum of the partner's archiving rules within the institution.

Integrity and confidentiality

Personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures.

All personal data will be handled with appropriate security measures applied. This means:

- *Data sets with personal data will be stored at a Google Drive server at the that complies with all GDPR regulations and is ISO 27001 certified.*
- *Access to this Google Drivel be managed by the project management and will be given only to people who need to access the data. Access can be retracted if necessary.*
- *All people with access to the personal data files will need to sign a confidentiality agreement.*

Accountability

The controller shall be responsible for, and be able to demonstrate compliance with the GDPR.

At project level, the project management is responsible for the correct data management within the project.

6. DocksTheFuture approach to privacy and data protection

On the basis of the abovementioned regulations, it is possible to define the following requirements in relation to privacy, data protection and security:

- **Minimisation:** DocksTheFuture must only handle minimal data (that is, the personal data that is effectively required for the conduction of the project) about participants.
- **Transparency:** the project will inform data subjects about which data will be stored, who these data will be transmitted to and for which purpose, and about locations in which data may be stored or processed.

- **Consent:** Consents have to be handled allowing the users to agree the transmission and storage of personal data. The consent text included Deliverable 7.1 must specify which data will be stored, who they will be transmitted to and for which purpose for the sake of transparency. An applicant, who does not provide this consent for data necessary for the participation process, will not be allowed to participate.
- **Purpose specification and limitation:** personal data must be collected just for the specified purposes of the participation process and not further processed in a way incompatible with those purposes. Moreover, DocksTheFuture partners must ensure that personal data are not (illegally) processed for further purposes. Thus, those participating in project activities have to receive a legal note specifying this matter.
- **Erasure of data:** personal data must be kept in a form that only allow forthe identification of data subjects for no longer than is strictly necessary for the purposes for which the data were collected or for which they are further processed. Personal data that are not necessary any more must be erased or truly anonymised.
- **Anonymity:** The DocksTheFuture consortium must ensure anonymity by applying two strategies. On the one hand, anonymity will be granted through data generalisation and; on the other hand, stakeholders' participation to the project will be anonymous except they voluntarily decide otherwise

The abovementioned requirements translate into three pillars:

1. **Confidentiality and anonymity** – Confidentiality will be guaranteed whenever possible. The only exemption can be in some cases for the project partners directly interacting with a group of participants (e.g., focus group). The Consortium will not make publicly accessible any personal data. Anonymity will be granted through generalisation.
2. **Informed consent** – The informed consent policy requires that each participant will provide his/her informed consent prior to the start of any activity involving him/her. All people involved in the project activities (interviews, focus groups, workshops) will be asked to read and sign an Informed Consent Form explaining how personal data will be collected, managed and stored.
3. **Circulation of the information limited to the minimum required** for processing and preparing the anonymous open data sets –The consortium will never pass on or publish the data without first protecting participants' identities. No irrelevant information will be collected; at all times, the gathering of private information will follow the principle of proportionality by which only the information strictly required to achieve the project objectives will be collected. In all cases, the right of data cancellation will allow all users to request the removal of their data at any time

7. FAIR (Findable, Accessible, Interoperable and Re-usable) Data within Docks The Future

DMP component Issues to be 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 generated/collected

- Specify if existing data is being re-used (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

The purpose of data collection in Docks The Future is understanding opinions and getting feedbacks on the Port of The Future of proper active stakeholders - defined as groups or organizations having an interest or concern in the project impacts namely individuals and organisations in order to collect their opinions and find out their views about the "Port of the Future" concepts, topics and projects. This will Include the consultation with the European Technological Platforms on transport sector (for example, Waterborne and ALICE), European innovation partnerships, JTIs, KICs.Consortium Members have (individually) a consolidated relevant selected Stakeholders list.

The following datasets are being collected:

- Notes and minutes of brainstorms and workshops and pictires of the events(.doc format, jpeg/png)
- Recordings and notes from interviews with stakeholders (.mp4, .doc format)
- Transcribed notes/recordings or otherwise 'cleaned up' or categorised data. (.doc, .xls format)

No data is being re-used. The data will be collected/generated before during, or after project meetings and through interviews with stakeholders.

The data will probably not exceed 2 GB, where the main part of the storage will be taken up by the recordings.

The data will be useful for other project partners and in the future for other research and innovation groups or organizations developing innovative ideas about ports.

2. 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

The following metadata will be created for the data files:

- Author
- Institutional affiliation
- Contact e-mail
- Alternative contact in the organizations
- Date of production
- Occasion of production

Further metadata might be added at the end of the project.

All data files will be named so as to reflect clearly their point of origin in the Docks The Future structure as well as their content. For instance, minutes data from the meeting with experts in work package 1 will be named “yyy mmm ddd DTF –WP1-meeting with experts”.

No further deviations from the intended FAIR principles are foreseen at this point.

3. 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
   - Specify how access will be provided in case there are any restrictions

Data will initially be closed to allow verification of its accuracy within the project. Once verified and published all data will be made openly available. Where possible raw data will be made available however some data requires additional processing and interpretation to make it accessible to a third party, in these cases the raw data will not be made available but we will make the processed results available. Data related to project events, workshops, webinars, etc will be made available on the docks the future website. No specific software tools to access the data are needed. No further deviations from the intended FAIR principles are foreseen at this point

4. Making data interoperable
   - Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.
   - Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

The collected data will be ordered so as to make clear the relationship between questions being asked and answers being given. It will also be clear to which category the different respondents belong (consortium members, external stakeholder). Data will be fully interoperable – a full unrestricted access will be provided to datasets that are stored in data files of standard data formats, compliant with almost all available software applications. No specific ontologies or vocabularies will be used for creation of metadata, thus allowing for an unrestricted and easy interdisciplinary use

5. Increase data re-use (through clarifying licences)
   - Specify how the data will be licenced to permit the widest reuse possible
   - Specify when the data will be made available for re-use. 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 re-use of some data is restricted, explain why
   - Describe data quality assurance processes
   - Specify the length of time for which the data will remain re-usable

Datasets will be publicly available. Information to be available at the later stage of the project. To be decided by owners/ partners of the datasets. It is not envisaged that Docks The Future will seek patents. The data collected, processed and analyzed during the project will be made openly available following deadlines (for deliverables as the datasets. All datasets are expected to be publicly available by the end of the project. The Docks The Future general rule will be that data produced after lifetime of the project will be useable by third parties. For shared information, standard format, proper documentation will guarantee re-usability by third parties. The data are expected to remain re-usable (and maintained by the partner/ owner) as long as possible after the project ended,

6. Allocation of resouces
   - Estimate the costs for making your data FAIR. Describe how you intend to cover these costs
   - Clearly identify responsibilities for data management in your project
   - Describe costs and potential value of long term preservation

Data will be stored at the coordinator's repository, and will be kept maintained, at least, for 5 years after the end of the project (with a possibility of further prolongation for extra years). Data management responsible will be the Project Coordinator (Circle). No additional costs will be made for the project management data.

7. Data Security
   - Address data recovery as well as secure storage and transfer of sensitive data

Circle maintains a backup archive of all data collected within the project. After the Docks The Future lifetime, the dataset will remain on Circle's server and will be managed by the coordinator.

8. 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

No legal or ethical issues that can have an impact on data sharing arise at the moment

8. Open Research Data Framework

The project is part of the Horizon2020 Open Research Data Pilot (ORD pilot) that "aims to make the research data generated by selected Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access. This implies that the DocksTheFuture Consortium will deposit data on which research findings are based and/or data with a long-term value. Furthermore, Open Research Data will allow other scholars to carry on studies, hence fostering the general impact of the project itself.

As the EC states, Research Data "refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation. [...] Users can normally access, mine, exploit, reproduce and disseminate openly accessible research data free of charge". However, the ORD pilot does not force the research teams to share all the data. There is in fact a constant need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IRP), privacy concerns, and security.

The DocksTheFuture consortium adopts the best practice the ORD pilot encourages – that is, "as open as possible, as closed as necessary". Given the legal framework for privacy and data protection, in what follows the strategy the Consortium adopts to manage data and to make them findable, accessible, interoperable and re-usable (F.A.I.R.) is presented.